Cub-H02970-V13-P003265

No. 1 (New Series No. 201) January 1981

ANALYSIS

Edited by
Christopher Kirwan

13

CONTENTS

tion of types and other types of quotation	Laurence Goldstein
initeness in identity	John Broome
ining Donnellan's distinction	Jeffrey King *and* Michael Liston
aratives	C. J. W. Williams
aratives and degrees	Adam Morton
nent	C. J. F. Williams
le worlds and diagonalization	A. W. Moore
in and the identity theory	Michael Lockwood
urious egocentricity	Jeremy Butterfield
ledge, reliable methods, and Nozick	David Gordon
bility: subjective and mathematical	P. J. R. Millican
ledge and non-accidental guessing	Terry Dartnall
nportance of being completely wrong	Roy A. Sorensen
for a 3-valued logic	Fred Johnson
ffects of principles and of actions	M. C. Geach

BASIL BLACKWELL PUBLISHER · OXFORD

ISSN 0003–2638

P3265

METAPHILOSOPHY

Editor Terrell Ward Bynum

Vol. 14, nos. 3 & 4 — July/October 1983

Symposium on Obedience and Morality	
Milgram's dilemma of obedience	MARIO F. MORELLI
Reply	STANLEY MILGRAM
Methods of ecological ethics	ROBIN ATTFIELD
Rights and social freedom	JOHN D. HARMAN
A defense of ethical relativism	ROBERT L. ARRINGTON
Life is not absurd	S. JACK ODELL

Also articles on Quine by Paul A. Roth, David J. Ross & Thomas E. Wartenberg, 'semantic ascent' by Sallas Willard, Feminist epistemology and women scientists by Alan Soble, Rouseau on reason, nature and women, by Genevieve Lloyd, Self-refuting arguments by Michael Stack.

1983 Subscription rates
Institutions: £25.00 (UK) £34.40 (Overseas) $57.00 (USA) $70.00 (Canada)
Individuals: £15.95 (UK) £19.50 (Overseas) $36.50 (USA) $45.00 (Canada)
Orders, with payment, to: Sue Dommett, Journals Department, Basil Blackwell, 108 Cowley Road, Oxford, OX4 1JF.

Journal for the Theory of Social Behaviour

General Editors: P. F. Secord, P. Collett,
and S. Moscovici

Volume 13 no. 3. — October 1983

On the Prospects for a Nomothetic Theory of Social Structure	DOUGLAS V. PORPORA
Replies to Porpora	
Comments on the Prospects for a Nomothetic Theory of Social Structure	PETER M. BLAU
Idiographic vs. Nomothetic Explanation	JONATHAN H. TURNER
Macro Versus Micro Interpretation	NORBERT WILEY
Causality, Historical Particularism and Other Errors in Sociological Discourse	BRUCE H. MAYHEW
Realism and Social Scientific Theory	JEFFREY ISAAC
Rejoinder	
	DOUGLAS V. PORPORA
Individual Papers	
Reduction, Epigenesis and Explanation	PETER T. MANICAS

Annual subscriptions: Institutions £25.50 (inland) £31.25 (overseas) $65.00 (USA) $80.00 (Canada);
Individuals £12.95 (inland) £15.95 (overseas) $35.95 (USA) $44.00 (Canada)

Orders to: JOURNALS DEPARTMENT, **BASIL BLACKWELL,**
108 Cowley Road, Oxford OX4 1JF, England.

Morality and Conflict

STUART HAMPSHIRE

Stuart Hampshire has had an influence far beyond the traditional bounds of academic philosophy. His lucid prose and compelling arguments have made his works essential reading for students and scholars alike, while his concern for practical issues in moral and philosophical philosophy has attracted a wide following among the general public. In these essays he explores the complex relationship between morality and conflict.
186 pages, £15.00 (0 631 13336 4)

Philosophy and Sociology of Science

An Introduction

STEWART RICHARDS

Modern science is inseparable from the society that supports it. Yet now, more than ever before, the scientist is encouraged to specialize and narrow his field of vision. Stewart Richards argues that it is vital to reverse this trend, and that all students of science – natural and social – should look at the context, roots and structure of their disciplines.
224 pages, hardback £17.50 (0 631 13411 5)
paperback £5.95 (0 631 13414 X)

Freud and Human Nature

ILHAM DILMAN

Interest in the writings of Freud seems never to decline, but it is rare that an important book appears which discusses Freud's belief that human nature exists independently from the society in which an individual develops. This book provides this: a thorough critical examination of the various issues raised by Freud's ideas of sexual and moral development.
216 pages, £15.00 (0 631 13373 9)

Kant's Political Philosophy

HOWARD WILLIAMS

Although highly regarded for his critical philosophy, Kant has rarely been thought of as a significant political theorist. In this comprehensive new account of Kant's political and social philosophy, Howard Williams attempts to redress this imbalance, and shows that Kant's political views are not only interesting in themselves but shed considerable light on the critical philosophy out of which they grew.
304 pages, £20.00 (0 631 13123 X)

Basil Blackwell

P 3265

Beyond Objectivism and Relativism

RICHARD J. BERNSTEIN

Philosophy is again questioning its goals and assumptions, a malaise signified by the growing opposition between objectivism and relativism, between rationality and irrationality, between realism and anti-realism. Contemporary thinking is becoming increasingly polarized and attempts to escape from this entire framework of analysis have all too frequently been assimilated to these standard oppositions. This major new book represents a courageous attempt to break out of this mould.
320 pages, hardback £17.50 (0 631 13429 8)
paperback £6.50 (0 631 13437 9)

Practical Reason

G.H. von WRIGHT

This first volume of Professor von Wright's **Philosophical Papers** explores realms which he has made peculiarly his own: the logic of action; and the logical study of normative concepts and discourse, known by the name he coined for it in 195 as 'deontic logic'. **Practical Reason** testifies to his unique contribution in applying the tool of logic to discourse about norms and actions and to the practical necessities on which normative concepts are founded.
224 pages, £15.00 (0 631 13315 1)

Philosophical Logic

G.H. von WRIGHT

Since the publication twenty-five years ago of **Logical Studies**, G.H. von Wright has pursued his odyssey through the territories of philosophical logic. This volume of his **Philosophical Papers** traces his path. Many of the papers included have hitherto been virtually inaccessible; the author has revised and in some cases expanded them specially for this book.
158 pages, £15.00 (0 631 13316 X)

Reason in Human Affairs

HERBERT SIMON

The capacity for using reason in making choices between alternative courses of action is seen as one of the defining characteristics of the human species. What can reason or thinking do for us and what can it not do? This is the important question addressed in this ambitious but accessible book. In it, Professor Simon examines the principal theoretical frameworks designed to describe and explain reasoning processes.
126 pages, £8.95 (0 631 13409 3)

Basil Blackwell

Philosophical Investigations

Edited by **D. Z. Phillips**

Philosophical Investigations publishes articles which discuss, extend or criticise philosophical approaches pioneered by J. L. Austin, Gilbert Ryle and Ludwig Wittgenstein.

Volume 7 No. 1 January 1984

Articles:

The Language of Sense Data and Private Experience – I
Notes taken by RUSH RHEES of Wittgenstein's Lectures, 1936
Wittgenstein On Private Languages: It Takes Two To Talk
BENJAMIN F. ARMSTRONG, Jr.

Critical Notices:

G. E. M. Anscombe, *Collected Philosophical Papers*
Vols. I – III RENFORD BAMBROUGH
K. R. Popper, *The Open Universe: An Argument for Indeterminism* and *Quantum Theory and the Schism in Physics*
ANTONY FLEW
John H. Whittaker, *Matters of Faith and Matters of Principle*
WALFORD GEALY

Published quarterly: January, April, July and October
Subscriptions Volume 7 (1984)
Institutions: *£25.00 (UK); £30.75 (Overseas); $58.50 (US); $71.70 (Canada)*
Individuals: *£11.95 (UK); £14.65 (Overseas); $28.00 (US); $34.30 (Canada)*

Orders, with remittance, or requests for specimen copies and further details should be sent to:
Sue Dommett, Journals Department, Basil Blackwell,
108 Cowley Road, Oxford, OX4 1JF, England.

Basil Blackwell · Oxford · England

QUOTATION OF TYPES AND OTHER TYPES OF QUOTATION

By LAURENCE GOLDSTEIN

THE type/token distinction, apparently a simple one, is beset by worrisome difficulties, and attempts to give an account of types have been plagued by confusion. A suggestion of Hugly and Sayward [11] that types are *classes* has been successfully rebutted, in my opinion, by Peter Simons [17], but Simons' preferred theory, a *pattern* approach to types, also seems unsatisfactory. In this note, I shall first register some complaints about Simons' theory, then present an alternative account of type-talk which, I claim, is simpler and more general in that the type/token distinction falls out as a special case.

Simons proposes the convention of using single quotation marks to name types, double quotation marks to name patterns, the Greek letters ξ, η as place-holders for names and φ, ψ as place-holders for predicate letters. A single expression may exemplify several different patterns. Thus, any token of '*Faa*' exemplifies all of the patterns "Faa", "$Fa\xi$", "$F\xi a$", "$F\xi\xi$", "$F\xi\eta$", "φaa", "$\varphi a\xi$", "$\varphi\xi a$", "$\varphi\xi\xi$", "$\varphi\xi\eta$". A clue to Simon's view of the relation between patterns and types is to be found in his assertion that, as a limiting case, a fully materially determined pattern is just a type: "*Faa*" = '*Faa*'.

Simons says that 'we cannot *isolate* a pattern from all material whatsoever'. He doesn't say why, but presumably the reason is that a pattern is a pattern *in* material, not a physical part *of* material and can be no more prised off material than its shape can be prised off a table. Patterns are thus erganzungsbedürftig – in need of completion, incapable of independent existence. According to Simons, patterns have *tokens* which are therefore dependent, incomplete particulars. Following Husserl, Simons calls such particulars 'moments'.

Simons asserts that his theory is nominalistically acceptable, but in order to see how dubious this claim is one need only look at the alleged relation between patterns and their moments. Here, surely, we have a prime example of the 'one over many'. Indeed, the relation is compared by Simons to that holding between many cricket shots (all dependent particulars) and one stroke, the cover drive, that they all exemplify. For Simons, a pattern is not a general word (which a nominalist could countenance), but a *universal* in which real (though dependent) particulars participate.

It seems that Simons is alive to this danger of raising nominalists' hackles. The avoidance manoeuvre he adopts, which is not without historical precedent, amounts to *particularizing* the putative universals. Like G. F. Stout [18], who sought to avoid the universal *redness* by positing numerically different particular rednesses – to each red particular its own token redness – Simons tries to escape universal patterns by positing particular tokens of patterns. However, whereas Stout's 'abstract particulars' are contained in ordinary objects in the straightforward mereological sense of 'contains' of which Simons approves, it is not clear whether this is the sense invoked when Simons claims that a token of '*Faa*' *contains* a pattern-token "$F\xi\eta$". Just how the token pattern is included in, and is dependent upon the token sign remains obscure.

One suggestion might be that token patterns are the result of mentally abstracting from token signs. That is to say, by selectively ignoring certain features of signs one may reveal some aspect common to all members of a set of signs, and one could then say that each member of the set contains a token of this aspect or pattern. Thus, some sense could be given to the claim that pattern-tokens are dependent particulars. They are dependent or incomplete in the sense that they are mind-dependent; they are mental entities contrived (abstracted) from an ontological base of token signs. This characterization of token patterns is not, however, available to Simons. For him, token patterns together with token signs *constitute* the ontological base from which types are abstracted (see [17], p. 202). His ontology, like Frege's, includes real incomplete particulars. But surely talk of real incomplete particulars is metaphorical, standing in need of explanation rather than standing *as* explanation of anything (see [3], pp. 71–82). Further, Simons' project of explaining, in a nominalistically acceptable way, recurrent features in terms of particularized universals is dogged by just those familiar difficulties that dog Stout's theory (see [1], pp. 77–87) and Simons gives no indication of how such obstacles might be surmounted.

The dispute just initiated could, no doubt, be pursued at length, but I think a more fundamental question ought to be raised about whether the problem being disputed is a genuine or a spurious one. The alleged problem concerns the nature of types, but does our discourse about types really give rise to such a problem? Quine, enthusing about Bentham's theory of paraphrasis, suggests that when we find a term convenient but ontologically embarrassing, it may be possible to define it contextually, and so continue to enjoy the services of the term while disclaiming its denotation ([16], p. 68). Yet 'type' and 'pattern' are nouns, and nouns generally denote, so it is plausible to suppose that our type- and pattern-talk is talk of objects called types and patterns. The supposition is apparently vindicated by our being able to name types and patterns by surrounding tokens with appropriate quotation marks, so that it is possible to say, for example, that "$F\xi a$" is a pattern.

This line of reasoning, though smooth, is circular. For merely enclosing a printed token between quotation marks cannot serve to name a type or a pattern unless there is a type or pattern to be named. But whether there are types or patterns whose nature it is for us to explain is precisely what has been called into question.

To say that two tokens are of the same type is to say that the two are typographically similar. In thus speaking of types, we are speaking of *relations* (in particular, of relations of *similarity*) between tokens. One reason why we should resist the claim that there are such things as types (and patterns) is that to assimilate relations to things exposes one to Bradley's regress. Frege treats relations – or, in his terminology, functions – as unsaturated things, but this springs from an acute recognition of the categorial distinction between functions and objects, linked with an incorrect doctrine of predicate *reference* (see [3], p. 74). As Wittgenstein pointed out ([20], 3.1432), a relational expression does not refer to a thing; the occurrence in a statement of a relational expression expresses how things are related (see also [14]). So to ask for an account of the ontological nature of types betrays the category confusion of asking about the constitution of *how things are related*.

We have suggested that talk of types can be paraphrased as talk of relations of similarity (degrees of equiformity) among tokens. In this way, we may dissolve the problem about the nature of types by contextually defining it away. Now Simons also wants to 'trade in' (as he puts it) problematic sentences about types for unproblematic sentences about tokens, but what prompts his trading is a certain puzzle about types that he constructs. Curiously, he does not show how the puzzle is to be solved, but sees that it may be bypassed by jettisoning some of the locutions employed in setting it up. It will be instructive to examine this puzzle with a view to understanding just what is defective about the goods of which Simons wishes to divest himself.

The puzzle is as follows. Suppose that we have a simple first-order language which includes the names '*a*' and the formula '*Faa*' (where the quotation expressions – the quotation marks together with what lies between them – name types, not tokens). It seems that the formula contains two occurrences of the name '*a*'. But, if it contains two tokens of the name, then the formula is an impossible type-token hybrid. Yet, if it contains type expressions, then, since types are unique individuals, it contains two type '*a*'s which are then presumably tokens of a higher, second-order type. By considering more complex formulae such as '*Faa* ⊃ *Faa*', which similar reasoning reveals to be either an impossible hybrid or to contain third-order type '*a*'s, we see that a vicious infinite regress of types is in prospect. Neither alternative is attractive.

In setting up his puzzle, Simons treats quotation marks as *naming* devices. Yet, since names are semantically unstructured, it is nonsense to say that one name *contains* other names. If '*Faa*' were a name, as Simons contends it is, it would no more contain the *name* '*a*' than the name 'Martha' contains the name 'Art'. There has been a tendency, especially since the publication of Tarski's classic paper on truth [19], to regard expressions as names, but, as a number of authors have conclusively shown, that view is inadequate (e.g. [5], [10]). The correct way to bypass Simons' puzzle is to see that quotation marks do not serve to produce proper names or any other sort of name.

In talking of types, quotation is indispensable, so a sound account of quotation should do much to explicate such talk. Davidson (op. cit.) defends a *demonstrative* theory according to which quotation marks may be read 'the expression with the shape pictured herein'. For spoken language, where normally there occur no explicit indicators of quotation (although 'finger-dance quotes' or double 'tch's are sometimes used) it is the phonetic shape that is demontrated by the (explicit or implicit) quotation marks. Now, this account, although certainly on the right track, will not do as it stands. For when Elvis says 'Baby, don't say "don't",' he is not just requiring his baby to refrain, when confronted with a certain request, from uttering tokens of the same phonetic shape as 'don't', but from uttering any tokens that *mean* the same. Davidson is correct in claiming that quotation marks serve as a pointing gesture, but it is a familiar fact that pure ostension may leave the intended reference underdetermined. One may be referring to the particular token contained within the quotation marks to any token that looks similar to the one exhibited, to any token that means the same, to any token that has the same syntax, i.e. when referring to phemes (see [2]), to any token that displays a similar pattern and so on. Although in practice misunderstandings rarely arise, the interests of perspicuity would be served by employing different styles of quotation marks (see [8], [9]). Thus, in the Elvis example, the meaning-quotes of Kaplan [12] would be appropriate for the second 'don't'.

In each of these cases, the quoted token is demonstrated, or pointed to, as a sample, and does not occur as a grammatical part of the sentence. The quotation marks are the *demonstrative* and the sample the *demonstratum*. (For related views on demonstration and on indirect quotation, see Kaplan [13] and Davidson [4].) One can think of the quotation marks as arrows pointing to a sample which could be remote from the sentence. This will remove the temptation of regarding the quotation expression as a name, and so remove the problem of discovering what it names. An alternative convention, simpler for the typesetter, uses 'this' and a colon as a demonstrative, and places the sample after the colon. For example,

(a1) There are four tokens of the type 'quotation' on this page.

becomes

(a2) There are four tokens on this page typographically similar to this: quotation.

Another example, but one in which the criteria of identity for the demonstratum are not typographical, is provided by Frege's difficulty about the concept *horse*. Following Frege, I here use italics for the same end (as he puts it) as quotation marks around 'horse'. In the phrase 'the concept *horse*' the quotation expression is construed by Frege as a name in apposition (see [7], p. 46 and especially the critique of Shoenflies [6], p. 177). Hence, according to his theory, it must refer to an *object*, not a concept, and we can make clear what sort of object it is by letting the singular term 'the concept *horse*' stand for it. This sits ill, of course, with the sharp distinction that Frege wishes to maintain between *concept* and *object*.

At first, Frege tends to play down the difficulty, ascribing it to an 'awkwardness of language' ([17], p. 46), a 'peculiar obstacle in the way of an understanding' which can nevertheless be overcome by relying on a charitable reader meeting him halfway, not begrudging a 'pinch of salt' ([7], p. 54). Later, however, he comes to treat the difficulty more seriously: 'If I want to speak of a concept, language, with an almost irresistible force, compels me to use an inappropriate expression which obscures – I might almost say falsifies – the thought' ([6], p. 119; see also p. 177).

On our modified Davidsonian account of quotation, Frege's difficulty vanishes, for we treat quotation not as a naming device, but as a *pointing* device, and in

(b1) The concept *horse* is a concept.

what is pointed to is a concept. Ostension of the concept is mediated by a written token, which token, in the context of (b1), has what scholastic writers termed '*suppositio simplex*'. The written token is pointed to *qua* concept, rather like when I point to a car and say 'I would like that for the walls of my room' (meaning the *colour* of the car). Quine's notion of deferred ostension ([15], p. 40) is a distant relative of the idea I'm employing here. As a paraphrase of (b1), the best I can offer is the highly inelegant, but in spirit Fregean,

(b2) In any judgement that would be overtly expressed as:
Such and such is a horse,
whatever in the judgement makes the same contribution to the sense of the whole as does:
is a horse
to the overt judgement is a concept.

Here the first ostension is to a token judgement, the second to a *predicative* phrase.

Hong Kong University,
Pokfulam Road,
Hong Kong

© Laurence Goldstein 1984

REFERENCES

[1] D. M. Armstrong, *Nominalism and Realism, Volume 1*, Cambridge: Cambridge University Press, 1978.

[2] J. L. Austin, *How To Do Things With Words*, Oxford: Oxford University Press, 1962.

[3] D. A. Bell, *Frege's Theory of Judgement*, Oxford: Clarendon Press, 1979.

[4] D. Davidson, 'On Saying That', *Synthese*, 19 (1968-9), 130-46.

[5] D. Davidson, 'Quotation', *Theory and Decision*, 11 (1979), 27-40.

[6] G. Frege, *Posthumous Writings*, Oxford: Basil Blackwell, 1979.

[7] P. Geach and M. Black (eds), *Translations from the Philosophical Writings of Gottlob Frege*, Oxford: Basil Blackwell, 1970.

[8] L. Goldstein, 'Categories of Linguistic Aspects and Grelling's Paradox', *Linguistics and Philosophy*, 4 (1981), 405-21.

[9] L. Goldstein, 'Linguistic Aspects, Meaninglessness and Paradox: A Rejoinder to John David Stone', *Linguistics and Philosophy*, 4 (1982), 575-92.

[10] S. Haack, 'Mentioning Expressions', *Logique et Analyse*, 67-8 (1975), 277-94.

[11] P. Hugly and C. Sayward, 'Expressions and Tokens', *Analysis*, 41.4 (October 1981), 181-7.

[12] D. Kaplan, 'Quantifying In', *Synthese*, 19 (1968-9), 178-214.

[13] D. Kaplan, 'Demonstratives, Draft #2', UCLA mimeograph, 1971.

[14] P. Long, 'On *Expressing* Functions and Properties', *Analysis*, 36.3 (March 1976), 118-25.

[15] W. V. Quine, *Ontological Relativity and Other Essays*, New York: Columbia University Press, 1969.

[16] W. V. Quine, 'Five Milestones of Empiricism' in his *Theories and Things*, Cambridge, Mass.: Harvard University Press, 1981.

[17] P. Simons, 'Token Resistance', *Analysis*, 42.4 (October 1982), 195-203.

[18] G. F. Stout, 'Are the Characteristics of Particular Things Universal or Particular?', *Proceedings of the Aristotelian Society*, Supp. Vol. 3 (1923), 114-22.

[19] A. Tarski, 'The Concept of Truth in Formalised Languages' in his *Logic, Semantics and Metamathematics*, Oxford: Clarendon Press, 1956.

[20] L. Wittgenstein, *Tractatus Logico-Philosophicus*, London: Routledge Kegan Paul, 1922.

INDEFINITENESS IN IDENTITY

By JOHN BROOME

A CLUB is constituted by its rules and society's conventions, and these may not be enough to determine everything about it. The rules and conventions, for instance, may not specify the procedure for removing a dishonest treasurer. And they may not determine precisely what counts as the club's demise. Consequently, circumstances can arise that make it unclear whether or not a club started at some date is the same as a club that exists at some later date. (Suppose, for example, that the club has no meetings for a long time and then some people, including perhaps a few of the original members, start to meet again under the same name.) The indefiniteness here is not merely epistemological. Even if we knew everything there is to be known about the case we might still not know whether the clubs are the same or different. The question of identity has no answer; the facts do not determine one. The club's constitution leaves this particular question unsettled.

This account of the matter – that it is indeterminate whether or not the club existing earlier is the same as the one existing later – seems perfectly transparent; nothing about it is hard to understand. But an alternative account is possible. We could say that the act of creating an object, such as a club, is incomplete unless the object is defined in enough detail to settle all questions of identity; unless its constitution has this much precision no club has been created. This account, though, is obscure. Few actual clubs can have such precise constitutions, so according to this account most people who think they belong to a club must actually not do so. Indeed they must not belong to anything, not even an uncompleted club, because the same unsettled questions of identity will arise about uncompleted clubs as arise about clubs. It is hard, then, to understand what exactly is supposed to be the condition of these people who think they belong to a club.

A third possible account insists that if this matter of identity appears to be indeterminate then the clubs must in fact be definitely different, or perhaps definitely the same. This too is obscure. For we can describe cases that plainly amount to the continued existence – or the revival – of the same club, and we can describe cases that plainly amount to the dying of one club and the creation of another, and by varying the conditions gradually we can arrive at cases that are intermediate between these two. It is hard to understand what could make a sharp division between them, as this account insists there must be.

So we have, at least, a good *prima facie* example of indefinite identity.

II

However, it has recently been argued by Gareth Evans ('Can there be vague objects?', ANALYSIS, 38.4, October 1978, p. 208) and Nathan Salmon (*Reference and Essence*, Blackwell 1982, pp. 245-6) that there can be no indefiniteness in identity. They would deny that what I called the transparent account of my example can be the right one. Both authors offer similar arguments, and since Evans's is the more fully stated I reproduce it here:

> Let 'a' and 'b' be singular terms such that the sentence '$a = b$' is of indeterminate truth value, and let us allow for the expression of the idea of indeterminacy by the sentential operator '∇'.
>
> Then we have:
>
> (1) $\nabla(a = b)$.
>
> (1) reports a fact about b which we may express by ascribing to it the property '$\hat{x}[\nabla(x = a)]$':
>
> (2) $\hat{x}[\nabla(x = a)]\, b$.

> But we have:
>
> (3) $\sim\nabla(a = a)$
>
> and hence:
>
> (4) $\sim\hat{x}[\nabla(x = a)]\,a$.
>
> But by Leibniz's Law, we may derive from (2) and (4):
>
> (5) $\sim(a = b)$
>
> contradicting the assumption, with which we began, that the identity statement '$a = b$' is of indeterminate truth value.

From Leibnitz's Law, to make it explicit, Evans is taking this premise:

(6) $(a = b) \supset (Ea \equiv Eb)$

where E is the property $\hat{x}[\nabla(x = a)]$.

Harold Noonan has shown (ANALYSIS, 42.1, January 1982, pp. 3-6) that this argument breaks down if either of the terms 'a' or 'b' does not definitely denote an object. He agrees with Evans that there can be no indefiniteness in the identity of *objects*, but points out (what Evans never denied) that an identity *statement* might still have no definite truth value because its terms might have no definite denotation. I, on the other hand, shall claim that Evans's argument is invalid even if each term accurately denotes a single object. So I shall assume throughout that it does.

Compare this argument, which purports to prove that it can never be indefinite whether or not someone is tall. Let 'n' be the name of a person such that the sentence 'n is tall' is of indeterminate truth value. Let us use '∇' as Evans does. Then we have:

(1′) ∇(n is tall).

(1′) reports a fact about n which we may express by ascribing to him the property $\hat{x}[\nabla(x \text{ is tall})]$:

(2′) $\hat{x}[\nabla(x \text{ is tall})]\,n$.

But we have (playing a role analogous to (6))

(6′) (n is tall) $\supset \sim$ Bn

where B is the property $\hat{x}[\nabla(x \text{ is tall})]$. From (2′) and (6′) we may derive:

(5′) $\sim$ (n is tall)

contradicting the assumption, with which we began, that the statement 'n is tall' is of indeterminate truth value.

This argument and Evans's are both invalid for the same reason. (6) and (6′) are conditional statements whose consequents we have proved false. It follows that their antecedents are not true. But it does not follow that their antecedents are false because they may be neither true nor false. And that is exactly what we originally assumed they were.

Section III formalizes this point in the case of Evans's argument.

III

We must allow for a third truth value besides true (t) and false (f), say undetermined (u). The logical operators that appear in the argument are, without further definition, ambiguous. Take negation for instance. Does the conclusion ~ (a = b) assert that 'a = b' is false or that it is not true, i.e. either false or undetermined? Evans evidently means the former since he takes ~ (a = b) to contradict that 'a = b' is undetermined, so let us define '~' by this, the most conventional, truth table:

p	~p
f	t
u	u
t	f

The connective '⊃' in (6) is, like '~', ambiguous. To define it we shall have to specify which combinations of truth values for 'a = b' and for 'Ea ≡ Eb' are allowed by (6) and which are not. And of course we must design our definition so that (6) really is an instance of Leibnitz's Law.

The ground of Leibnitz's Law is that if a is identical to b then a and b are one object, and an object has a property if and only if it has it. This licenses us to say only that

(7) If 'a = b' is true then 'Ea ≡ Eb' is true.

In other words it rules out that 'a = b' should be true and 'Ea ≡ Eb' false, and also that 'a = b' should be true and 'Ea ≡ Eb' undetermined, but no other combination. So if (6) is to express Leibnitz's Law the truth table of '⊃' must be

	p	q	p ⊃ q
1	f	f	t
2	f	u	t
3	f	t	t
4	u	f	t
5	u	u	t
6	u	t	t
7	t	f	f
8	t	u	f
9	t	t	t

(Perhaps there should be a u instead of the f in line 8, but nothing turns on this.)

According to line 4 of this table, (6) does not rule out the possibility that 'a = b' might be undetermined and 'Ea ≡ Eb' false. But if this is a possibility Evans's proof fails. Leibnitz's Law, represented by (6), will not allow us to derive his (5) from (2) and (4). All we can derive from (2) and (4) is that 'a = b' is either false or undetermined. We cannot derive that it is false.

If the proof is to be valid our table must have f or at most u instead of t in row 4. But is there any reason to put anything other than t here? As well as (7) does Leibnitz's Law also give us

(8) If 'a = b' is undetermined then 'Ea ≡ Eb' is not false?

I can see no reason to think so. It may, perhaps, seem rather surprising that as 'a = b' drops, as it were, from being true to being undetermined, 'Ea ≡ Eb' should be allowed to drop all the way from being true to being false. But that should not be surprising if we remember that the property E (like B in my parallel argument) is specially selected to bring about exactly that. E is the property that x has when 'a = x' is undetermined. It is a good objection to (8) to point out that when 'a = b' is undetermined then b has this property and a does not, so that 'Ea ≡ Eb' is actually false. Certainly it is begging the question to assume (8) in order to prove that 'a = b' can never be undetermined. For against this conclusion we have a good *prima facie* example, whereas in favour of (8) we have no argument at all.

IV

I have been assuming that Evans wishes to oppose the view that 'a = b' might be neither true nor false. That seems plain enough from the passage I quoted where he speaks of 'a = b' as being 'of indeterminate truth value'. Since we are not concerned with epistemological indeterminateness, he can only mean that it is neither true nor false. And Salmon describes the view he opposes as 'there is a pair of entities ... such that it is vague (neither true nor false, indeterminate, there is no objective fact of the matter) whether they are one and the same thing'.

However, since both authors rely on classical two-value logic it is conceivable that they really have in mind a different opponent who is ready to concede to them that 'a = b' must be either true or false. (And Evans at one point does hint at a possible non-truth-functional interpretation for his '∇'.) So it is worth asking whether their argument is effective against this more amenable opponent.

The difficulty, of course, is in knowing what he can be supposed to be saying. The least he can believe is that 'a = b' is neither definitely true nor definitely false. To reconcile this with a belief that '$a = b$' must be either true or false we shall have to make a difference between a proposition's being true and its being definitely true. Let us write 'p is definitely true' as 'Δp' (so $\Delta p \supset p$ but not: $p \supset \Delta p$). Then the claim that it is indefinite whether or not $a = b$ must be:

$$(9) \quad \sim \Delta(a = b) \text{ and} \sim \Delta \sim (a = b)$$

and this must be equivalent to Evans's $\nabla(a = b)$.

But under this interpretation Evans's argument achieves no contradiction.[1] He derives

$$\sim (a = b),$$

which is consistent with (9). It may seem that by showing a is not the same as b Evans has dislodged his opponent from the position he would like to occupy, namely that a is neither definitely b nor definitely not b. But evidently not. If (9) states the position he wants to occupy then he can still occupy it. And if (9) does not state his position what does?

The only position Evans has made untenable is

$$\nabla(a = b) \text{ and } (a = b)$$

or equivalently

$$\sim \Delta(a = b) \text{ and } (a = b),$$

i.e. that a is b but a is not definitely b. Now some believers in indefinite identity might actually take a view something like this (as David Wiggins pointed out to me). They might think that '$a = b$' can have degrees of truth, so that it could be true to some degree without being definitely or completely true. This might be the right view to take in cases like the club where there is a range of possibilities running from definite identity at one end to definite non-identity at the other, with indefinite identity in between. Evans's argument could be taken to show to those who take this view that if '$a = b$' has any degree of truth less than definite truth then it cannot be true at all.

To those who hold a similar opinion about tallness my corresponding argument seems to show that if 'a is tall' has any degree of truth less than definite truth then it cannot be true at all. But under this interpretation my corresponding argument fails because its premise

$$(6') \quad (n \text{ is tall}) \supset \sim Bn$$

---

[1] At the end of his paper Evans derives the conclusion $\Delta \sim (a = b)$ from the initial supposition $\nabla(a = b)$. This is undoubtedly a flat contradiction. But his derivation assumes (6) as a premise and it works only if the operator 'Δ' generates a logic as strong as S5 (with 'definitely' replacing 'necessarily' in the standard interpretation). So it is effective only against someone who thinks that $(a = b)$ is either true or false but neither definitely true nor definitely false, and who attaches a sense to 'definitely' that keeps (6) true and also generates a logic as strong as S5. I do not know of such a person. The 'definitely' that appears in a degrees-of-truth analysis, for instance, makes (6) false, as I shall be showing.

is false if it is understood to say that the consequent is true to some degree whenever the antecedent is true to some degree. The property B is $\hat{x}\,[\nabla(x \text{ is tall})]$, and '$\nabla(x \text{ is tall})$' must now be read as

> 'x is tall' has a degree of truth other than definite truth or definite falsity.

So if 'n is tall' is true to some degree but not definitely true, n will definitely have property B. Then the antecedent in (6′) will be true to some degree but the consequent definitely false. Of course, whenever the antecedent is definitely true then so is the consequent. This shows that we cannot reason from

> If P is definitely true then Q is definitely true

to

> If P is true to some degree then Q is true to some degree.

Have we any better reason to rely on

(6) $(a = b) \supset (Ea \equiv Eb)$

understood as saying that

(10) If '$a = b$' is true to some degree then '$Ea \equiv Eb$' is true to some degree?

I think not. Certainly Leibnitz's Law tells us that

(11) If '$a = b$' is definitely true then '$Ea \equiv Eb$' is definitely true.

But I have just shown that we cannot infer (10) from (11). Against (10) there is this to be said. E is the property $\hat{x}\,[\nabla(a = x)]$, and '$\nabla(a = x)$' must now be read as

> '$a = x$' has a degree of truth other than definite truth or definite falsity.

So if '$a = b$' is true to some degree but not definitely true then b definitely has the property E and a definitely does not. Consequently '$Ea \equiv Eb$' is definitely false. To insist on (10) in order to prove that '$a = b$' cannot be true to some degree without being definitely true is once more begging the question.

V

I think we may continue to believe that identity can be indefinite.[2]

Department of Economics,
University of Bristol, Bristol BS8 1HY

© JOHN BROOME 1984

---

[2] I wish to thank: Derek Parfit for valuable comments and the example of the club, Christopher Peacocke for valuable comments, David Wiggins for his great patience in arguing with me, which very much helped me to understand the problems, the Social Sciences Research Council for supporting me while I worked on this paper and All Souls College, Oxford, for its hospitality during that time.

EXPLAINING DONNELLAN'S DISTINCTION

By JEFFREY KING *and* MICHAEL LISTON

ACCORDING to D. E. Over ('Effective and non-effective reference', ANALYSIS 43.2, March 1983, pp. 85-91), Donnellan's distinction between referential and attributive uses of definite descriptions is best explained in terms of the notion of an effective method. The core of Over's proposed explanation seems to be as follows:

> (E) If S uses '(ix)Fx' referentially, then S has (or associates with '(ix)Fx') an effective method for deciding which given object is the referent of '(ix)Fx'; if S has no such effective method, then S uses '(ix)Fx' attributively.

Over feels the need for some such explanation because he contends that Donnellan's examples involving uses of 'Smith's murderer' and other definite descriptions, although they illustrate a linguistic phenomenon, fail to explain it. While we are sympathetic to this complaint, we will argue that Over's explanation is no improvement.

To better explain Over's account and to motivate our own discussion, we will begin by considering an example from mathematics, where the intuitive notion of an effective method can be precisely captured in a number of ways. A speaker, S, using the sentence

> (1) The thirty-seventh prime is odd

might or might not have an effective method for picking out the thirty-seventh prime. He might, for example, have a primitive recursive function, f(n), which for an argument n calculates the nth prime. His knowledge of the truth of (1) might then be based on his having computed the value *a* of the function for the argument 37. Having noted that *a* is odd, S might use (1) to say of *a* that it is odd. On the other hand, S might know of no effective method for picking out the thirty-seventh prime, yet he might have knowledge that (1) is true – based, for example, on his general knowledge that all primes except two are odd. In the former case, it seems to be in keeping with Donnellan's distinction to say that S uses the description 'the thirty-seventh prime' referentially, in the latter, that he uses the description attributively. However, S might have the effective method f(n) yet use 'the thirty-seventh prime' attributively. Out of laziness he may prefer to appeal to his general knowledge rather than to calculate the value *a* (the particular number *a* not being important for the purposes at hand). It seems reasonable, then, not to equate referential use with possession of an effective method and attributive use with lack thereof, but instead to adopt with Over the weaker conditionals (E). According to Over, (E) has a number of advantages over Donnellan's own explanations.

Attractive though this account may be, it does not yield the main advantage Over claims for it: the explanation of 'what it is for a speaker (and his audience) to be able to pick out whom (or what) he is talking about' (p. 86). Consider the following example.[1] Suppose S, himself being slow mathematically and not knowing that all primes greater than two are odd, asserts (1) solely on the basis of having been given the number *a*, which he notes is odd, by Ramanujan, a recognized mathematical genius. Suppose also that in asserting (1), S uses the description 'the thirty-seventh prime' to pick out *a* and to say of it that it is odd. As such, his use is a paradigm of Donnellan's referential use. But though S can now pick out the thirty-seventh prime, in what sense does S have an *effective method* for picking out the thirty-seventh prime? Certainly S does not possess any analogue of our recursive function, f(n). Indeed S does not seem to be in possession of anything analogous to an effective method in the narrow sense of this term used in mathematics. Whatever S's ability to pick out *a* rests upon, it does not rest upon a mechanical method which, when applied correctly, is logically bound to produce the correct answer in a finite number of steps.

These considerations raise serious problems with (E). If (E) is to illuminate Donnellan's distinction and the notion of picking out a referent, the operative notion of effective method ought to be at least as clear as what it is intended to explicate and sufficiently general to capture Donnellan's everyday examples. We need either an account of the effective methods which purportedly operate in such examples as that given above or an explanation of how such methods are analogous to those used in mathematics. But given that Over does not provide the former and that the latter is likely to be unsuccessful for lack of an analogy, his explanation of S's referential use of 'the thirty-seventh prime' and S's ability to pick out its referent by appeal to the notion of effective method leaves us in the dark.

Our difficulty with (E) then is not that the generalized notion of effective method is unexplained.[2] Rather, given that it is unexplained and that in paradigm cases of referential use a speaker's ability to pick out a referent does not seem to be analogous to possession of an effective method in the narrow, mathematical sense, Over's explanation of Donnellan's distinction is no improvement on Donnellan's own account. Over has merely exchanged one label for another: he is still illustrating or describing a linguistic phenomenon, not explaining it.

© JEFFREY KING *and* MICHAEL LISTON 1984
University of California at San Diego,
La Jolla, California 92093, U.S.A.

---

[1] This is a modification of an example used by Over to illustrate the generalized notion of effective method. See his 'Predicative and Constructive Knowledge', *Analysis*, 42.3, June 1982, p. 143.

[2] Over himself mentions some of the difficulties with capturing precisely this generalized notion (op. cit. p. 145); however, they may be more serious than he realizes.

COMPARATIVES

By C. J. F. WILLIAMS

BARRY MILLER (ANALYSIS 43.1, January 1983, p. 18) disputes my contention (ANALYSIS 42.4, October 1982, p. 189) that I could not acquire the property of being shorter than my brother after ceasing to exist. He suggests that there is nothing paradoxical about saying 'Tom has grown to be taller than Bill' several years after Bill has died. He infers from this that there is nothing paradoxical in saying 'It has come to pass that Tom is taller than Bill', nor, accordingly, in saying 'It has come to pass that Bill is shorter than Tom'. In reply I should wish to point out that 'Tom has grown to be taller than Bill' in the context described is elliptical for 'Tom has grown to be taller than Bill was'. The converse of this is 'Bill was shorter than Tom has grown to be'; and if we insist on translating this into the 'come to pass' idiom, we shall get, not 'It has come to pass that Bill is shorter than Tom', which is false, nor 'It has come to pass that Bill was shorter than Tom', which is nonsense, but 'Bill was shorter than it has come to pass that Tom is', which is merely clumsy. There is no suggestion here that Bill has acquired a property.

The point is not a trivial one, since it allows attention to be drawn to a fact about comparatives which is not always appreciated. When asked to supply an example of a relational expression, a two-place predicable, we frequently produce '— is taller than . . .'. The choice is not a happy one. In 'Tom is taller than he was' the expression '— is taller than . .' is certainly not being used as a complete two-place predicate. In 'Tom is taller than Bill was' it is '— is taller than . . . was' that is the two-place predicate; and this shows clearly that '— is taller than . . .' in 'Tom is taller than Bill' is only a two-place predicate if it is understood as elliptical for '— is taller than . . . is'. The words 'is taller than' in so far as they occur univocally in 'Tom is taller than he was' and 'Tom is taller than Bill' are at most a fragment of a variety of predicables – of an irreflexive predicable in the case of '— is taller than . . . is', of a non-reflexive one in the case of '— is taller than . . . was'.

I find it difficult to give a syntactical analysis of the proposition 'Tom is taller than he was'. The two-place predicable '— is taller than . . . was' can be discerned in it, of course, and this can be seen to be subjected to the reflection operator (described by Professor Geach in *Reference and Generality*, §§82-4) to yield the one place predicable '(—; x, y) x is taller than y was' which is then attached to the proper name 'Tom'. But what of the fine structure of '— is taller than . . . was'? This is clearly a complex expression formed out of tense operators and some other expressions. These cannot include the one-place predicable '— is tall', since we should rightly regard '— is taller than . . .' as more fundamental than '— is tall' –

'___ is tall' being an obvious ellipsis for '___ is taller than the average', or the like. But '___ is taller than . . .' can only be interpreted as a predicable if it is taken as elliptical for '___ is taller than . . . is', and it is difficult to see how this predicable could occur in '___ is taller than . . . was'. If we took '___ is taller than . . . was' as the result of subjecting the second 'is' in '___ is taller than . . . is' to the past tense operator 'it was the case that ___' (as equivalent to '___ is taller than it was the case that . . . is'), we should have to distinguish this from 'It was the case that ___ is taller than . . . is', just as 'Bill was shorter than it has come to pass that Tom is' has to be distinguished from 'It has come to pass that Bill is shorter than Tom'. But how can there be two different ways of attaching a tense operator to a predicable? Traditional grammar counts 'than' as a conjunction, and if we were able as logicians to place it in the same syntactical category as, say, '___ or . . .', which also counts as a conjunction, we could analyse '___ is taller than it was the case that . . . is' in the same way as '___ is white or it was the case that . . . is square'. But '___ is square' is already a complete predicable, whereas '___ is' is merely the copula, a word which normally has no function other than that of converting a noun or an adjective into a verb phrase. What is the second 'is' doing in '___ is taller than . . . is'? I do not know the answer to this question, or indeed to the whole question about the correct analysis of propositions involving comparatives, and would like to use the pages of ANALYSIS to appeal for suggestions. One such suggestion appears on the following pages.

University of Bristol,
Queens Road, Bristol BS8 1RJ

© C. J. F. WILLIAMS 1984

COMPARATIVES AND DEGREES

By ADAM MORTON

I. DEGREES

THERE are many sentences in which the comparative form of the adjective cannot be taken as expressing a simple relation between the objects compared. For example

(a) Roger is fatter than we thought possible.
(b) Tony is fatter than he was yesterday.
(c) Tom has grown to be fatter than Bill.
(d) John ran faster than the Olympic record.
(e) Dennis drank more than the legal limit.

(b) and (c) present Christopher Williams' main worry, but the problem is rather more general. It is: how can we explain the logical form of comparatives in a way which makes (a) to (e) and the familiar 'Tom is fatter than Bill' all instances of the same construction?

Here is a possible solution. Sentences like (d) and (e) above are revealing because 'the Olympic record' names a speed (a distance plus a time) and 'the legal limit' names a quantity of alcohol. In some cases, then, we can construe a comparative as relating a subject to a *degree* to which it does or does not instantiate a predicate. Thus (d) and (e) may be construed as

(d′) $\exists s$ (John ran at speed s and s exceeded the Olympic record)
(e′) $\exists p$ $\exists q$ (Dennis drank p pints and it is illegal to drink more than q pints and p is greater than q)

This strategy will work in general. (a) and (b) can be taken as

(a′) $\exists g$ $\exists h$ (Roger's girth was g and we did not think it possible that Roger's girth was greater than h and g exceeds h)
(b′) $\exists g$ $\exists h$ (Tony's girth is now g and Tony's girth was h yesterday and g is greater than h)

(b′) shows how cross-temporal comparisons are possible.

Comparatives do appeal to two-place relations as thus construed, but not two-place relations between the things to which the corresponding non-comparative adjectives apply. The relation is between degrees. It is interesting to note that when the underlying non-comparative is two-place, the relation between degrees in the comparative is still only two-place, although there is a definite four-place appearance to the surface structure. Thus

(f) Adam loves Beth more than Carol loves David.

is represented as

(f′) $\exists d$ $\exists e$ (Adam loves Beth to degree d and Carol loves David to degree e and d is greater than e).

II. PARTIAL ORDERINGS

This last example should provoke a very natural worry. What *are* degrees and when is one greater than another? Love certainly does not come in *numerical* degrees, and it is for that matter far from obvious that fatness does either.

The relation between greater and lesser degrees of a predicate is not entirely unconstrained, though. It is evident that, for example:

x is F-er than y $\supset$ y is not F-er than x
x is not F-er than x
(x is F-er than y and y is F-er than z) $\supset$ x is F-er than z.

These do not entail that degrees correspond to numbers. Their force can be quite easily summed up, though, by the assumption that degrees form partial orderings. That is, we may assume that the degrees of any predicate are ordered by a relation that is reflexive, antisymmetric (i.e. Rxy and Ryx imply x = y) and transitive. An abstract partial ordering can be taken as the pattern common to a set of partially ordered domains which can be out into an order-preserving one-to-one correspondence.[1] The importance of an abstract partial ordering for our purposes is that it is not affected by the accidental nonexistence of elements having a particular predicate to a particular degree. There is a great variety of (abstract) partial orderings. In most of them there are incomparable elements, elements none of which bears *R* to any other. The representation of degrees by elements of partial orderings is thus compatible with the obvious possibility that *x* is *F* to degree *d* and *y* is *F* to degree *e*, and neither *d* nor *e* exceeds nor equals the other.

So, to put these pieces together, let me introduce just one more piece of apparatus, in the form of a function $\mathcal{L}$ which correlates with each predicate *F* its characteristic partial ordering $\mathcal{L}$ (F). Then the writing-out of '*a* is F-er than *b* was' is

> $\exists d\, \exists e\, \exists t$ (*a* is F to degree *d* now and *b* is F to degree *e* at time *t* and *d* exceeds *e* in $\mathcal{L}$ (F) and *t* is earlier than now)

and (e) 'John ran faster than the Olympic record' is

> $\exists d$ (John ran fast to degree *d* and the Olympic record is *e* and *d* exceeds *e* in $\mathcal{L}$ (running fast)).

It should be clear how all of (a) to (e) and Williams' examples are dealt with in these terms.

III. WHAT PARTIAL ORDERINGS?

The virtue of this approach is that it provides a uniform reading of the logical form of sentences involving comparative adjectives, which validates many of the more obvious inferences between them. In this connection it is worth noting that other constructions related to comparatives can be dealt with along the same lines. '*a* is just as F as *b*', '*a* is very F', and '*a* is too F to be G' are obviously related constructions, and can be handled in an obviously related way. They are, respectively

> $\exists d$ (*a* is F to degree *d* and *B* is F to degree *d*)
> $\exists d$ (*a* is F to degree *d* and *d* is high in $\mathcal{L}(F)$)

---

[1] For precise formulations and technical points see S. Maclane and G. Birkhoff *Algebra* (New York, Macmillan, 1968), Chapter 14, or, more likely to be on a philosopher's shelf, J. L. Bell and A. B. Slomson *Models and Ultraproducts* (North Holland, 1969), Chapter 1.

$\exists d$ (*a* is F to degree *d* and Necessarily (x) (e) (x is G and x is F to degree e) $\supset d$ exceeds e in $\mathcal{L}$(F))

Inferences between these constructions and comparatives (e.g. '*a* is too F to be G, *b* is G, therefore *a* is more F than *b*') are easily captured by this regimentation. At one time I thought that in order to capture some other facts about comparative degrees ought to conform to something more restricted than partial orderings, in particular that one ought to require degrees to be members of lattices. A lattice is a partial ordering in which for any two elements there is a greatest and a least of all the elements above and below both of them (see the references in footnote 1). But this suggestion seems to me now to be ruled out by the following example.

Imagine that in a society there is both an ecclesiastical and a temporal hierarchy, and that they are independent: there is no convention about how the two kinds of rank are compared and there is no top position common to both. Then 'is of high rank' and 'ranks higher than' both have application, but the partial ordering characterizing them is such that if *a* is of high ecclesiastical rank and *b* of high temporal rank then there will be no *c* such that *c* outranks both. And we may imagine that the institutions of the society forbid the joint possession of high ecclesiastical and temporal rank, so that it is part of the meaning of 'is of high rank' that this be so. Then $\mathcal{L}$ (is of high rank) is not a lattice. And I do not see how the fact that 'is of high rank' can be analysed as a disjunction can prevent it from being a proper predicate with a proper comparative.

If for any partial ordering Ω we can form a vague, but intelligible, predicate 'ranks highly in Ω', then we can repeat the argument of the last paragraph for this ordering. Then no restrictions could be imposed on the concept of a degree except the fairly mild one I have already postulated, that degrees form partial orderings. But I do not in fact think that this is the case. I think that we should exclude some partial orderings, for example one in which there are just two elements, neither one bearing the ordering relation to the other. Such an ordering could not be construed as that of the degrees of a predicate, because it does not allow for the fact that objects falling under the predicate must have that predicate to a greater degree than objects not falling under it.

Perhaps we can simply require that the partial ordering associated with a predicate allow a division into two classes, all of one bearing the ordering relation to all of the other. This would account for the obvious validity of

(Fx and $\sim$ Fy) $\supset$ x is F-er than y.

I am not convinced, though, either that this is the right way to handle this validity or that it represents the only reasonable further constraint on the partial orderings which represent the degrees of predicates. And I suggest that the interesting question we are left with, the successor to Williams' original question, is that of what the right set of constraints (expressed axiomatically or algebraically) on a partial ordering is, in order for it to be interpretable as the set of degrees of an intelligible predicate.

University of Bristol,
Queens Road, Bristol BS8 1RJ

© ADAM MORTON 1984

COMMENT

By C. J. F. WILLIAMS

THE price of Adam Morton's answer to my question may seem excessive, when we consider such simple propositions as 'Sarah has grown' or 'The radiator is heating up'. These propositions are equivalent to 'Sarah is taller than she was' and 'The radiator is hotter than it was'. Morton's analysis of these will then have the form

> $\exists d \, \exists e \, \exists t$ (a was F to degree d at time t and a is F to degree e now and t is earlier than now and e exceeds d in $\mathcal{L}$(F))

I might even want to expand this to

> $\exists d \, \exists e \, \exists t \, (a; x, y)$ (x was F to degree d at time t and y is F to degree e now and t is earlier than now and e exceeds d in $\mathcal{L}$(F))

Can growth and warming up be as complicated as this?

One reaction must be dismissed quickly. It will not do to say that growth and every sort of change which involves quantity, whether extensive or intensive, belong to a class of phenomena which cannot be further analysed. We cannot acquiesce in the notion that growth is, as Locke would put it, a simple idea. Whether we say that Sarah has grown or that Sarah overtops Richard we are either way employing the concept *taller than*; and the connections between the two-place predicable '— is taller than . . . is' and the one place predicable '— has grown' have to be traced. The cost may be the talk of degrees and partial orderings which Morton suggests, and if this is what it costs, we must pay up. But I am still open to lower bids.

University of Bristol,
Queens Road, Bristol BS8 1RJ

© C. J. F. WILLIAMS 1984

POSSIBLE WORLDS AND DIAGONALIZATION

By A. W. MOORE

IN Appendix 9 to 'Meaning, Quantification, Necessity' (Routledge and Kegan Paul, 1981), Martin Davies outlines an argument which he describes as an 'apparent paradox'. Interesting consequences arise from construing the argument not as a paradox but as a straightforward *reductio ad absurdum* of one of its premises, namely that 'for each proposition it is possible that X should have been thinking a thought at t whose content would have been specifiable by a sentence expressing that proposition'. If the argument is spelt out in detail, using the technique of diagonalization which lies at the heart of Cantor's theorem, then, on one view, it shows that there is a proposition which acts as a counterexample to this premise. On that view, the argument can be recast as follows.

Consider some man X and some time t. Let us abbreviate 'X is thinking a thought at t whose content would be specifiable by a sentence expressing the proposition p' by 'X is thinking the proposition p at t'. Note that this is to be understood as meaning that p is the *unique* proposition which X is thinking at t, for that is the only plausible interpretation under which Davies' original argument goes through. Now let f be an arbitrary function taking propositions as arguments, such that, for any proposition p, $f(p)$ is a possible world in which X is thinking p (alone) at t, if there is such a world, and is undefined otherwise. (We are guaranteed the existence of such a function if we assume the Axiom of Choice.) Then there is a proposition p^* for which f is indeed undefined, and which X therefore cannot be thinking at t. It is that proposition which is (or is determined by) the set of possible worlds w, such that, for some p, $f(p) = w$ and $w \notin p$. The function f is undefined for p^* as argument, because, if it weren't, there would be some possible world, say w^*, such that $f(p^*) = w^*$, and (since w^* cannot be $f(p)$ for any other p) it would follow that $w^* \in p^*$ if and only if $w^* \notin p^*$, a clear contradiction.

On this way of construing it, then, the argument shows that there is at least one proposition which X cannot be thinking at t – a proposition in each of whose member-worlds X will be thinking some proposition at t to which that world does not belong (a proposition, in other words, in each of whose member-worlds X will be thinking some proposition at t which is false in that world). This will be a proposition which entails (or is a subset of) the proposition that X is thinking some proposition at t which is false. But then arguably, the conclusion begins to lose its air of paradox; it is not too counterintuitive to say that there is a proposition which entails that the unique proposition that X is thinking at t is false, and which cannot itself be the unique proposition that X is thinking at t. What is particularly striking, however, is what happens

P 3265

when we turn our attention away from the property of being thought by X at t, and focus instead on the property of being the unique proposition expressed by the sentence S. A parallel argument shows that there is a proposition which entails that what S uniquely expresses is false, and which cannot itself be the unique proposition expressed by S. The significance of this is that what was perhaps an antecedently familiar conclusion arising from various semantic paradoxes has now been directly attained by the technique of diagonalization in its essentially set-theoretical guise. It is well known that the same technique led to the discovery of Russell's paradox: where we have exploited an argument showing that not every subset of the set of all possible worlds can be assigned a unique possible world, Russell exploited an argument showing that not every subset of the set of all sets can be assigned a unique set (not even itself). And although there is good reason to doubt the coherence of all this talk about possible worlds and propositions (partly, indeed, as Davies suggests, because of this very argument), this is surely one more interesting point of connection between the so-called set-theoretical and semantic paradoxes.[1]

University College, Oxford © A. W. MOORE 1984

---

[1] For various helpful comments, I am indebted to Dr S. R. Blamey, J. J. Campbell and an anonymous referee.

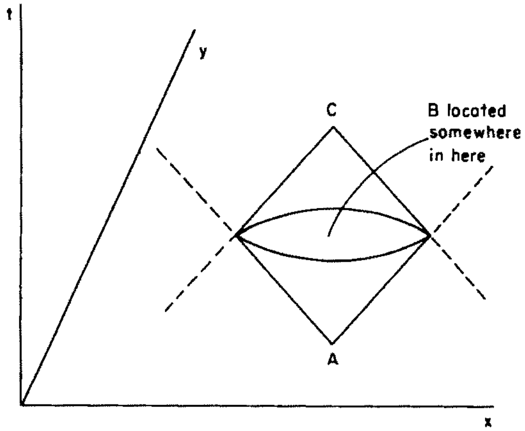
EINSTEIN AND THE IDENTITY THEORY

By MICHAEL LOCKWOOD

IN this note I shall attempt to scotch what to many people has seemed to be a serious objection to the theory that mental states are identical with certain events in the brain. Advocates of the identity theory maintain that the identification of mental states with brain states is no different in principle from other such identifications that we should all accept as scientifically compelling: a favourite example is the identification of flashes of lightning with electrical discharges. The question, then, is whether the proposed identification of mental states with brain states really is, in relevant respects, like that between lightning and electrical discharges, say, or whether there may not be some crucial disanalogy. Assume, for the sake of argument, that it will prove possible to establish the requisite correlations between mental events and events in the brain, so that the cases are alike to that extent.

One supposed disanalogy, of which much has been made in the philosophical literature, lies in the alleged impossibility of assigning a spatial location to a mental event – or more precisely, of doing so independently of some prior assumption that it was identical with such and such an event in the brain. A flash of lightning and an electrical discharge, so it has been argued, as well as being assigned times of occurrence, may be independently – that is, without assuming identity – assigned spatial locations. If, then, the assigned time *and the assigned spatial location* both turn out to coincide, as of course they would in this case, we have a very strong motive for identifying them: as Ayer has neatly put it, in conversation, 'they are competing for the same space'. But no such argument, it is alleged, is available in the case of mental states and brain states: the only grounds on which we could assign a spatial location to a mental state would be some prior decision to identify it with some physiological event the location of which was already known. There is no way in which we can first assign a spatial location to a given mental event and then use this as an argument for identifying it with whatever physiological happening turns out to have been going on at that place at the time in question. Indeed, given that this independent assignment of a spatial location to a mental event seems to be an impossibility *in principle*, some philosophers go so far as to claim that to speak of such events as having a spatial location – of my twinge of anxiety, say, as occurring seven centimetres behind my left eye – is actually meaningless.

This argument I believe to be mistaken. It is, I suggest, possible in principle to assign a spatial location to a mental event, independently of any prior identification with some physical event. That is to say, it can be done if we make certain assumptions. Assume (1) that the special theory of relativity is correct. Now, according to special relativity, any two events that are temporally separated, with respect to one frame of reference, must be spatially separated with respect to some other frame. So if we assume (2) that mental events are located in physical time, that they belong to the same temporal order as physical events, it follows immediately, given special relativity, that they are also in physical space, since events could hardly be spatially separated without being spatially located. That may not seem very helpful, since it does not as yet give us a *way* of spatially locating mental events; it just says that they must be located somewhere or other. But if we make a few more assumptions, we shall have a method of location. Assume (3), along with common sense, that physical events may cause mental ones, as when I sit on the proverbial pin and it punctures my skin, thereby causing pain, and (4) that mental events may cause physical ones, as when my pain causes me to wince or my embarrassment causes me to blush. These last two assumptions may seem too obvious to be worth stating, but of course there is nothing so obvious that a philosopher cannot be found to deny it. Finally (5), assume what is known technically, amongst physicists, as *causality*: the principle that causal influences cannot propagate faster than the speed of light. (Contrary to popular belief, this does not follow from special relativity alone; it is perfectly possible to postulate faster than light particles, *tachyons*, whose behaviour is entirely in accord with the Lorenz transformation. But there are also excellent reasons for not believing them to exist in reality.) This gives us all we need to assign at any rate an approximate spatial location to some mental events. Let A be a physical event that causes a mental event, B, which in turn causes a physical event, C. If we know the time of occurrence and spatial locations of A and C, then we can at least place bounds upon the time of occurrence *and spatial location* of the mental event B: it must lie within the intersection of the forward light cone of A and the backward light cone of C. A space-time diagram will make this clear. (The third spatial axis, the z axis, has been suppressed.)



If it were now possible to find pairs of physical events, standing to B as cause and effect respectively, that were separated by ever smaller temporal intervals, then it would be possible to define the spatial location of B with any required degree of precision.[1] This is a very strong supposition, of course, but even so I see no reason why someone who was disposed even to believe the correlational assumption behind the identity theory should be inclined to reject it. Most causal processes that we know of are continuous. (I ignore the inevitable setting in, at a certain point, of quantum indeterminacy.)

---

[1] Possible, that is, in principle: I am not suggesting that this is a remotely *practical* method of assigning to mental events spatial locations precise enough to be of use to the neurophysiologist.

I have put things this way, because of possible difficulties surrounding the notion of assigning a precise time of occurrence to a mental event, without correlating it with a physical one. Strictly, of course, it would suffice to find, for a mental event whose spatial location we were interested in, a physical cause *or* a physical effect – providing, that is, we had an independent fix on when the mental event occurred. For then the mental event would be known to be no further from the physical cause or effect than light could have travelled in the temporal interval separating them.

At this point in a philosophical paper, it is customary for the author to anticipate and attempt to counter certain possible lines of objection. But for my part I cannot think of any that could have any real appeal. I suppose someone might conceivably wish to reject the assumption that mental events are temporally related to physical events in the same way that physical events are temporally related to each other. But to deny that would be to wreak havoc with the notion that mental and physical events can be causally related to each other. It seems impossible to deny that if mental events are to be part of a causal order that includes physical events, they must belong to the same temporal order. The burden of this paper has been to show that, if Einstein was right about space and time, they must belong to the same spatial order as well.[2]

Department for External Studies,
Oxford University,
Wellington Square, Oxford OX1 2JA

© MICHAEL LOCKWOOD 1984

---

[2] Robert Weingard, of Rutgers University, remarked to me several years ago that if mental events are in time then, according to relativity, they must be in space as well. He did not elaborate, but that remark was the ultimate inspiration for this paper.

ON SPURIOUS EGOCENTRICITY

By JEREMY BUTTERFIELD

MY aim is to refute an argument in Prior's paper, 'On Spurious Egocentricity', that purports to show that tensed sentences cannot be analysed token-reflexively. I admit that Prior has some other, good, arguments to this end, both in this paper and elsewhere (Butterfield 1984, Section 2). But I think it is worth exposing this argument; it is intriguing, and raises some interesting points about belief as well as time.

The token-reflexive analyses proposed by Reichenbach (1948, pp. 284-7) and Smart (1949, p. 492) suppress verbal tenses and other temporal indexicals like 'now' and 'five days ago'; instead they use tenseless verbs and convey temporally indexical information by stating whether the fact the sentence reports is earlier than, later than, or contemporaneous with 'this utterance'. Prior draws an analogy between the 'egocentric theory of truth' that claims that 'It is true that . . .' and 'I believe that . . .' have the same meaning; and the token-reflexive theory that identifies 'It is now the case that . . .' and 'It is contemporaneous with this utterance that . . .'. (He also draws an analogy between Ramsey's redundancy theory of truth and his own redundancy theory of the present; but we do not need to consider this more positive part of his argument.) Prior allows that the egocentric theory of truth and the token-reflexive theory have an analogous attraction: one can sincerely say 'It is true that p' and 'I believe that p' in just the same circumstances (1968, p. 15); similarly for 'It is now the case that p' and 'It is contemporaneous with this utterance that p' (1968, p. 19).

But, he claims, they also have two analogous defects. The first is that 'there must surely be many truths which no one believes'; and similarly 'surely things like the solidifying of the earth were... present... without there being any utterances for them to be contemporaneous with' (1968, p. 20). I accept that this *is* a defect. That is to say, here is a good argument against the egocentric and token-reflexive theories (Butterfield 1984, Section 2). But I reject Prior's suggestion (1968, pp. 15-16, 19) that there is a second analogous defect.

Prior states this second defect, for the egocentric theory, by considering the following argument (taken from the exercises of Whateley's *Elements of Logic*):

> He who believes himself to be always in the right in his beliefs, claims infallibility: you always believe yourself to be right in your beliefs: therefore you claim infallibility.

He says that this argument is invalid since 'always believing oneself to be in the right' is not the same as 'believing oneself always to be in the right right'. We all do the former – we all believe our beliefs to be true. But we do not do the latter. However, says Prior, if the egocentric theory of truth were correct, the argument would be valid. If that theory were correct, 'in the right' would mean 'believed by me' and then we would all believe ourselves always to be in the right (in addition to always believing ourselves to be in the right). The fact that the argument is invalid of course refutes the egocentric theory.

Similarly, Prior claims (1968, p. 19), if 'It is now the case that...' and 'It is contemporaneous with this utterance that...' meant the same, the following argument would be valid:

> What is always contemporaneous with whatever is happening, must go on forever; but always my utterances are contemporaneous with whatever is happening; therefore I never stop talking.

Since the argument is not valid, these expressions do not mean the same.

This argument of Prior's is obscure. It is not clear what he takes the form of Whateley's argument to be; nor how the egocentric theory of truth is supposed to licence the inference from 'always believing oneself to be in the right' to 'believing oneself always to be in the right'. *Prima facie*, these phrases would be symbolized as: '$(p)B(Bp \rightarrow Tp)$' and '$B((p)(Bp \rightarrow Tp))$' respectively. Here 'B' stands for 'One believes that' and 'T' for 'It is true that'; 'always' is interpreted as a universal quantifier over propositions, which is quite a common usage. The invalidity of Whateley's argument can then apparently be blamed on inverting the order of the quantifier '(p)' and 'B' – which certainly matches the phrases' surface structure. And if one accepts that '(p)' and 'T' do commute, one can explain why the argument would be valid if one could identify 'B' and 'T': one would be validly inferring $T((p)(Bp \rightarrow Tp))$ from $(p)T(Bp \rightarrow Tp)$.

However, the situation is not so simple. The trouble is that under some natural assumptions, '(p)' commutes with 'B' just as it does with 'T'. If 'B' is treated like '$\Box$' in modal logic, so that $\ulcorner B\alpha \urcorner$ is true at a world iff α is true at all the doxastically alternative worlds, and if the range of the quantifier is the same at all worlds (say, some set of sets of worlds), then for any formula α, $\ulcorner B(p)\alpha \equiv (p)B\alpha \urcorner$ will be valid. (In the meta-language the formula corresponds to inverting the order of two universal quantifiers, one over worlds and the other over sets of worlds.)

Of course, it is questionable whether '(p)' should have the same range at all worlds; for example, perhaps a proposition referring to an object *a* should not exist in worlds where *a* does not exist (Prior 1957, p. 34f.; Fine 1977). And if we give '(p)' different ranges at different worlds, it will commute with T (given the obvious truth-clause: $\ulcorner T\alpha \urcorner$ is true at a world iff α is) but not with B. But clearly the invalidity of Whateley's argument should not turn on subtleties like this. This is brought home by the following example. Suppose my beliefs are accurate and exhaustive about what objects the world contains: that is, each of my doxastic alternatives contains just the actual objects. Then even if we tie a proposition's existing in a world to the existence there of the mentioned objects, the range of '(p)' will be the same at my doxastic alternatives, and '(p)' will commute with 'B'. But in this situation I surely should not claim infallibility: I may have false beliefs about the properties of objects (so that the actual world is not one of my doxastic alternatives).

So can we show that Whateley's argument is invalid, and that it would be valid if we could identify 'B' and 'T', without appealing only to a quantifier shift? Yes. We symbolize 'believing oneself always to be in the right', i.e. believing oneself infallible, as before: '$B((p)(Bp \rightarrow Tp))$'. But we symbolize 'always believing oneself to be in the right', in the sense in which we all do this just because belief aims at truth, by using a 'B' in the consequent: '$(p)(Bp \rightarrow BTp)$'. (We could use '$\equiv$' in place of '$\rightarrow$' without affecting the subsequent argument.) There is no reason to think this premiss implies the infallibility claim '$B((p)(Bp \rightarrow Tp))$'; it doesn't on the elementary semantics above. Nor do we get an implication if we add a 'B' to the premiss to give either '$B(p)(Bp \rightarrow BTp)$' or '$(p)B(Bp \rightarrow BTp)$'. On the other hand, if we identify B and T, we can argue as follows: $(p)(Bp \rightarrow BTp)$; so $(p)(Bp \rightarrow TTp)$; so $(p)(Bp \rightarrow Tp)$; so $T(p)(Bp \rightarrow Tp)$; so $B(p)(Bp \rightarrow Tp)$. Similarly for the variant premisses with an extra 'B'.

So what about the argument to the effect that I never stop talking? We should distinguish two questions. First, is this argument, which Prior presents as an analogue of Whateley's argument, in fact an analogue? Second, does the analogue of Whateley's argument provide an argument against the token-reflexive approach? I shall consider these in turn.

I believe the answer to the first question is No. It is clear first of all that this argument turns upon an ambiguity in 'always'. For the first (major) premiss to be true, 'always' must mean 'at all times'. But for the second (minor) premiss to be true, 'always' must mean 'at all times at which I speak'; that is, 'always' is here used as a restricted universal quantifier over times. Such a usage is quite common; when we say 'Socrates always lived in Athens', 'always' means 'at all times at which Socrates existed' not 'at all times' (Lewis 1975, p. 4; Stirling 1982, Chapter 6). Similarly, when we say 'Socrates is necessarily human', 'necessarily' means 'in all worlds in which Socrates exists' not 'in all worlds' (Kripke 1971, p. 137; Wiggins 1976, pp. 290, 301). If 'always' had the same range in the two premisses, the argument certainly would be valid.

I see no analogous ambiguity in a quantifier in Whateley's argument, on which its invalidity could be blamed. Furthermore, I see no way of formalizing the argument to the effect that I never stop talking, using say 'S' for 'It is now the case that...' and 'C' for 'It is contemporaneous with this utterance that...', in such a way that it gets the same form as Whateley's argument, once we follow Prior's basic analogy and replace 'S' by 'T' and 'C' by 'B'. For example, the second (minor) premiss might be formalized either as '$(p)(Sp \rightarrow Cp)$' (which is true) or, to capture the 'always', as '$A((p)(Sp \rightarrow Cp))$', where 'A' is read as 'It is at all times true that...' and 'C' refers to an utterance spoken at whatever time we are carried to by A. (The second formalization is false since there are times with no utterances.) But either formalization is a far cry from the analogue of the true second premiss of Whateley's argument, viz. $(p)(Cp \rightarrow CSp)$.

But I shall not give more details on this score, since it is more important to address our second question: does the real analogue of Whateley's argument — whether or not Prior supplies it — give an argument against the token-reflexive approach?

The answer is No. To do so, the argument must be invalid, but rendered valid by identifying ‘C’ and ‘S’. On my analysis, Whateley's argument infers the infallibility claim $B((p)(Bp \rightarrow Tp))$ from the claim that belief aims at truth, $(p)(Bp \rightarrow BTp)$. So the analogue of Whateley's argument must infer ‘$C((p)(Cp \rightarrow Sp))$’ from ‘$(p)(Cp \rightarrow CSp)$’. But such an inference may well be valid; after all its conclusion is arguably true, since ‘$(p)(Cp \rightarrow Sp)$’ is arguably true, as is its premiss ‘$(p)(Cp \rightarrow CSp)$’. So the fact that identifying ‘C’ and ‘S’ makes the inference valid (by a chain exactly analogous to that given above in the paragraph beginning ‘So can we show’) is no argument against that identification. I conclude that Whateley's argument provides no ammunition against the token-reflexive theory.

A final point. One can of course object to my analysis of Whateley's argument; perhaps another analysis would really have an invalid analogue which the token-reflexive theory would have to count valid. But I doubt it. For one thing, such an analogy would have to skirt around a clear difference between the two cases. This is that we want ‘$p \equiv Sp$’ and ‘$p \equiv Cp$’ to be valid for all present-tensed p; but though we want ‘$p \equiv Tp$’ to be valid for all true p, we do not want ‘$p \equiv Bp$’ (or either conjunct) to be valid for all true p.

Jesus College,
Cambridge CB5 8BL

© JEREMY BUTTERFIELD 1984

REFERENCES

J. Butterfield, ‘Indexicals and Tense’, in *Exercises in Analysis*, ed. I. Hacking, Cambridge: C.U.P., 1984.

K. Fine, ‘Properties, Propositions and Sets’, *Journal of Philosophical Logic*, 6 (1977), 135-91.

S. Kripke, ‘Identity and Necessity’, in *Identity and Individuation*, ed. M. K. Munitz, New York: N.Y.U., 1971.

D. Lewis, ‘Always’, in *Formal Semantics of Natural Language*, ed. E. Kennan, Cambridge: C.U.P., 1975.

A. Prior, ‘On Spurious Egocentricity’, in *Papers on Time and Tense*, Oxford: Clarendon, 1968, reprinted from *Philosophy*, 42 (1967), 326-35.

H. Reichenbach, *The Elements of Symbolic Logic*, New York: Macmillan, 1948.

J. Smart, ‘The River of Time’, *Mind*, 58 (1949), 483-94.

C. Stirling, *The Foundations of Logical Analyses of Tense*, York University PhD Thesis, 1982.

D. Wiggins, The *De Re* Must, in *Truth and Meaning*, ed. G. Evans and J. McDowell, Oxford: Clarendon, 1976.

KNOWLEDGE, RELIABLE METHODS, AND NOZICK

By DAVID GORDON

IN ANALYSIS 43.4 (October 1983) B. J. Garrett rejects a counterexample, proposed by Raymond Martin, to Nozick's analysis of the conditions of knowledge. As Garrett points out, Martin's example fails to satisfy Nozick's condition (3), unless 'the possible worlds in which Gumshoe does not win the first and Tagalong wins the second race are sufficiently "far" from the actual world as to be irrelevant to the assessment of S's putative claim to know that Gumshoe won the first' (p. 182). Since reasons have not been given why these possible worlds should be considered irrelevant, we have no reason to think that condition (3) is satisfied; and Martin's case is therefore no counterexample to Nozick.

I think that Garrett is right, if *in fact* the possible worlds he mentions are not too far out. But, Martin might reply, why need he give *reasons* for thinking these possible worlds far out – suppose they just are far out. And instances can easily be supplied in which this will be true. For example, suppose Tagalong is a wretched horse who really doesn't belong on the track. There will be no possible worlds close to the actual world in which Tagalong wins the second race and, obviously, none in which Gumshoe does not win the first and Tagalong wins the second. In this case, condition (3) is satisfied. Is, then, Martin's case a counterexample to Nozick under these circumstances?

I do not see that it is. All of Nozick's conditions are indeed met; but why is Martin's example not then a case in which S does know that Gumshoe won the first race? S is using a reliable method in the sense that his method obtains the right answer, in all possible worlds close to the actual world, to his question, 'did Gumshoe win the first race?'

A reliable method, in Nozick's use, is one that will probably give correct answers (*Philosophical Explanations*, pp. 264–7). Obviously, the reliability of a method depends on how the reference class is specified, and if S were to apply the method he uses to other horses in other races, he would no doubt quickly come to grief. But why assume, which was not given in Martin's example, that S intends his method to apply elsewhere?

It may be true that S does not have good grounds for thinking his method reliable: if he bets on Tagalong in the second race, he presumably does not think Tagalong an impossibly bad horse. Why is it a requirement for knowledge, however, that one have good grounds to believe one's method reliable?

Unless one adopts such a requirement, Garrett's counter examples do not succeed. In his first example, X believes that A's father is a philosopher on the basis of being told that B's father is a philosopher. A and B are brothers, and, on the essentialist assumption of this example, necessarily so. On Nozick's analysis, X knows that A's father is a philosopher, even though he does not know that A and B are brothers. Garrett objects that 'this is surely false – it was just good luck that, on this occasion, X's method happened to be applied to two brothers' (p. 3). But X's method will, if brothers are necessarily brothers, always obtain the right answer to his question, viz., whether A's father is a philosopher. Why does it count against its being knowledge that if X were to apply the same method to two unrelated people C and D, he would frequently find himself in error? X is using a reliable method, applied to brothers, whether or not he realizes that this method cannot be used for unrelated persons. On Garrett's essentialist view, it is precisely *not* a matter of luck that the method works: it couldn't have come out otherwise. Once more, why is it a requirement of knowledge that one have good grounds for thinking one's method reliable?

Garrett's counter here is obvious. He might say that the reference class by which the reliability of the method is to be judged is the class of all pairs of persons, whether or not related. With this reference class the method is unreliable. Although it seems to me plausible to say here that X is then using an unreliable method with a reliable submethod, though he does not know this, let us disregard this point. Assuming that X's method is unreliable, is Garrett's case a counterexample to Nozick?

It will certainly be true that X does not have a justified belief that A's father is a philosopher. This is a separate point from whether one knows one's method is reliable. On Nozick's 'externalistic' account, if one's method is reliable, one's belief is justified, even if one doesn't know that the method is reliable. But Nozick, although thinking it plausible that having a justified belief is a necessary condition for knowledge, prefers to leave this question unsettled (*Philosophical Explanations*, p. 267). His list of conditions is *not* an attempt to solve the 'problem of the fourth condition', i.e. to add whatever is needed to justified true belief to solve Gettier cases and other problems. His conditions are a *replacement* for the justified true belief analysis.[1] In the absence of argument that justified true belief is a necessary condition for knowledge, then, even if Garrett can show a case in which one can meet Nozick's conditions while using an unreliable method, he won't have arrived at a clear counterexample to Nozick.

Garrett's second example seems to me also open to objection. In this one Y believes, correctly, that if the number of people in Paris is even, it is the sum of two odd primes. Y's belief is correct because, for the example, we suppose that necessarily any even number is the sum of two odd primes. But Y does not know this. Then, Garrett claims, 'via his method, Y does not know the contingent truth that the number of people in Paris is the sum of two odd primes because he does not know that every large even number is the sum of two primes' (p. 183). But even if one strengthens the example by assuming that, apart from cases involving the residents of Paris, Y has no beliefs at all about whether even numbers are the sums of odd primes, Garrett has not provided a reason to reject Y's claim to knowledge. In the example, Y would not know why his method works: but it necessarily must work. Why is this not enough?

---

[1] Incidentally, Nozick's four conditions can be reduced to three. Since the possible worlds close to the actual world in applying condition (4) include the actual world, condition (2) is redundant.

Garrett's final example challenges Nozick's thesis when applied to necessary truths, in which case condition (3) does not apply. Garrett imagines that Z believes Fermat's Last Theorem (for the purposes of this example we suppose the theorem to be necessarily true) because no counterexamples have been discovered to it. Z satisfied conditions (1), (2) and (4) and, on Nozick's analysis, knows that Fermat's Last Theorem is true. But Garrett counters, 'Z does not know that Fermat's Last Theorem is true. His belief to that effect was reached via an unreliable method without the aid of any reliable sub-method' (p. 183).

In what way is Z's method unreliable? Is it that Z will, if he believes *any* mathematical proposition to which no counterexamples have been discovered, believe many false propositions, except on the highly unlikely assumption that no mathematical propositions not now known to have counterexamples will ever have them? But whether this situation obtains will depend on which mathematical propositions Z considers candidates for his method. If he believes true only theorems to which mathematicians have failed for a very prolonged period to find counterexamples, his method does not seem unreliable.[2] If, on the other hand, he applies his method more indiscriminately, Z will be using an unreliable method. But, then, as before, why should one not say that his unreliable method has a reliable submethod – believing mathematical propositions to which no counterexamples have been discovered will work when applied to necessary truths. Z will not know when his method is reliable; but, once more, why must he? Further, as with Garrett's first example, even if the method is unreliable, it is a counterexample only if justified belief is required for knowledge.

The heart of the difference between Nozick and Garrett emerges in Garrett's concluding paragraph. He suggests that Nozick has the task of resolving 'the problems which arise for his theory from certain situations in which a belief that ϕ is based upon a belief that ψ (where the believer is unaware of the logical or conceptual relation which holds between ϕ and ψ) ...' (p. 5). On Nozick's analysis, it isn't necessary to be aware of such logical or conceptual relations in order to know that ϕ. These relations explain *why* one's belief is true: but why is knowing this needed for having a method that works? In the absence of such a requirement, a supporter of Nozick can easily accommodate Garrett's examples as genuine cases of knowledge.[3]

---

[2] So far as the reliability of Z's method is concerned, Professor Nozick has pointed out to me in conversation that one way we in fact find out which propositions are necessary is by discovering those for which we cannot devise counterexamples. Proofs ultimately depend on axioms, which are accepted in the way just indicated.

6238 Orange Street,
Los Angeles, California 90048, U.S.A.

© David Gordon 1984

---

[3] I am grateful to Professor Nozick for many discussions about his theory of knowledge. He should not, however, be held responsible for anything in this article.

PROBABILITY: SUBJECTIVE AND MATHEMATICAL

By P. J. R. Millican

In 'Is Epistemic Preferability Transitive?' (Analysis 41.3, June 1981), Roy Sorensen attempted to give a counterexample to show that the 'more likely to be correct than' relation is not always transitive. My refutation of this argument ('On the Transitivity of Epistemic Preferability', Analysis 42.2, March 1982) focused on Sorensen's erroneous assumption that iterated probabilities may be collapsed, that if p implies *probably q*, and q implies *probably r*, so that p implies *probably (probably r)*, then p must imply *probably r tout court*. Nowhere do I suggest that such an assumption might be justified even under some alternative, non-mathematical interpretation of 'probable', and yet in his reply ('Subjective Probability and Indifference', Analysis, 43.1, January 1983) Sorensen appears to take for granted that this is the case. He writes: 'Peter Millican has pointed out that my argument ... is unsound *if* epistemic preferability is given a probabilistic interpretation', and he then goes on to 'provide *another* counterexample ... that *also* shows that epistemic preferability cannot be given a probabilistic interpretation' (my italics). He thus appears to enlist me to the ranks of those who *reject* the probabilistic view, despite the fact that I use probabilistic reasoning precisely to highlight 'the flaw in Sorensen's argument'! This misunderstanding I should like to rectify.

Let us suppose that a company director wishes to recruit two new employees but, because of his difficulty in choosing amongst the ten equally qualified applicants, decides to eliminate them one by one until he is left with two to fill the posts. As one of these ten, I am likely (epistemically) to survive the first elimination (since only one of the ten will fail to do so), and having become one of the remaining nine, I am also likely (epistemically) to survive the second. Similar reasoning would seem to show that I am likely to survive each elimination in turn, and if iterated likelihoods are allowed to collapse, this has the consequence that I am likely to be appointed. When there are five times as many applicants as jobs, such a result is surely absurd under any interpretation of 'likely', epistemic or otherwise!

Without his implausible collapsing of iterated probabilities, Sorensen's purported counterexample to the transitivity of epistemic preferability has absolutely no force whatever. In his later article, however, he presents a second apparent counterexample (strictly, to the transitivity of epistemic *indifference*), which we must now consider:

> Suppose there are 21 buckets of water $b_1, b_2, \ldots, b_{21}$ such that bucket b_i contains water at i degrees centigrade. Jim makes pairwise comparisons between b_i and b_{i+1} where $1 \leqslant i < 21$ by putting his hands in the appropriate buckets. Since Jim cannot discern a temperature difference of 1 degree by this method, he is epistemically indifferent between propositions p_i and p_{i+1} where 'p_i' reads 'The bucket containing water of the highest temperature of all 21 buckets is bucket i'. So if the epistemic indifference relation is transitive, Jim should be indifferent between p_1 and p_{21}. However, Jim can discern a temperature difference of 20 degrees and so he is not indifferent between p_1 and p_{21}. Therefore, epistemic indifference is not transitive.

This argument contains a subtle but devastating flaw, which is disguised by Sorensen's use of the present tense to describe all of Jim's actions and judgements. For transitivity has been undermined only if Jim is epistemically indifferent between p_i and p_{i+1} *at the same time* as he epistemically prefers p_{21} to p_1, and once the story is filled out a little, it becomes far from clear that this could be the case. Presumably Jim's preference for p_{21} over p_1 comes either from an explicit comparison, or else from a gradual realization that the temperature of the buckets has increased. Either way, this new information provides him with additional data which might require a revision of his previous judgements of epistemic indifference between p_i and p_{i+1}. Sorensen's case is obviously incomplete until he has dismissed such a possibility, and this he cannot do, for it can in fact be shown that once Jim knows b_{21} to be hotter than b_1, he does indeed have reason to prefer p_{i+1} to p_i in every case ($1 \leqslant i < 21$).

Let us suppose that Jim is able to detect a difference in temperature if and only if that difference is at least D degrees (such a sensory threshold will not of course be so well-defined in practice, but this makes no substantial difference to the argument). Thus he is faced with a series of indiscernible differences $d_1, d_2, d_3, \ldots, d_{20}$ (where d_i is the increase in temperature between b_i and b_{i+1}), all of them between $-D$ and $+D$, which together add up to an overall rise in temperature of D or more. Now *before comparing* b_1 *with* b_{21}, and in the absence of any further information about these indiscernible differences, Jim is quite epistemically indifferent between the possible values of, say, d_1 in the range $-D$ to $+D$: he knows that d_1 lies somewhere within this range, but he does not know where. The same is true of d_2, d_3 etc.; so as far as he is concerned, for all i, the value of d_i has a uniform probability density function over the range $-D$ to $+D$. Under these circumstances the probability that $d_i = 0$ is negligible, d_i having a 0.5 probability of being positive and 0.5 of being negative. It should be obvious that the same comment applies to the sum $S = d_1 + d_2 + d_3 + \ldots + d_{20}$, although this probability density function, over the range $-20D$ to $+20D$, is not uniform.[1] One important corollary is that, *before comparing* b_1 *with* b_{21}, the probability that $S \geqslant +D$ is less than 0.5 (the actual figure is approximately 0.349). In other words, it is unlikely that the difference in temperature between b_1 and b_{21} will be both positive and discernible.

There are 21 possibilities $h_0, h_1, h_2, \ldots, h_{20}$ concerning the direction of the 20 indiscernible differences $d_1 \ldots d_{20}$, where h_n represents the hypothesis that n of them are positive (representing a rise in temperature) and $(20-n)$ negative (representing a fall). Since each of these differences is independent of the others and equally as likely to be positive as negative, it follows that, *before comparing* b_1 *with* b_{21}, the probability that h_n is true yields a (binomial) distribution function in which h_{10} is the most likely hypothesis, while h_9 and h_{11}, h_8 and h_{12}, h_7 and $h_{13}, \ldots, h_0$ and h_{20}, are pairs of equally likely hypotheses, the pairs being themselves in order of decreasing probability. That h_n and h_{20-n} are indeed equally likely can be seen easily when the equal probability of positive and negative differences is combined with the observation that h_n postulates n positive and $(20-n)$ negative differences, while h_{20-n} postulates n negative and $(20-n)$ positive differences.

If h_n is true, then n of the 20 differences are positive and $(20-n)$ negative. So in the absence of any information to distinguish amongst them, any one of them stands an $n/20$ chance of being positive, and a $(20-n)/20$ chance of being negative. It should therefore be clear that *if, after comparing* b_1 *with* b_{21}, h_n *can be shown to be more probable for high values of* n *than for low values, then it will follow that any particular difference* d_i *is more likely to be positive than to be negative.* In this case, Sorensen's argument against the transitivity of epistemic indifference would be completely undermined, since p_{i+1} would be in all cases epistemically preferable to p_i, whereas Sorensen claims them to be epistemically indifferent. We now proceed to demonstrate, by an inverse probability argument using Bayes' Theorem, that *after comparing* b_1

---

[1] By the Central Limit Theorem, the distribution will be approximately normal, with mean zero and variance $20D^2/3$. In fact, the argument in the text in no way depends upon a *uniform* distribution function even for d_i: the important result is simply that Jim has no reason to advance to suppose b_i to be hotter than b_{i+1} – something that Sorensen's argument itself presupposes.

with b_{21}, Jim has reason to prefer h_{11} to h_9, h_{12} to h_8, h_{13} to h_7, and so on, thus fulfilling the antecedent of the italicized conditional.

Let us start by assuming that either h_9 or h_{11} is correct: we choose this pair because, as shown previously, the probabilities of h_9 and h_{11} 'initially' (i.e. before the comparison of b_1 and b_{21}) are exactly the same. Both hypotheses postulate nine positive indiscernible differences and nine negative – they differ in that according to h_9 the remaining two differences are negative, while according to h_{11} they are positive. It should be manifestly clear that the sum S of all twenty differences is more likely to exceed $+D$ if h_{11} is true than if h_9 is true.[2]

Bayes' Theorem states that the probability of an hypothesis H after observing some evidence E is equal to the initial probability of (H and E) divided by the initial probability of E. The initial probability of (H and E) is equal to the initial probability of H multiplied by the conditional probability that E is true given that H is true. Putting these together:

$$\text{Probability (H given E)}$$
$$= \frac{\text{Initial Probability of H} \times \text{Probability (E given H)}}{\text{Initial Probability of E}},$$

which is proportional to the conditional probability of E given H. Thus where we are given a single piece of evidence for two rival hypotheses which have the same initial probability, we should prefer whichever hypothesis makes that evidence more likely to have occurred, since this will give the greater value for the conditional probability of E.

In the case under discussion, we are comparing the hypotheses h_9 and h_{11} in the light of the 'evidence' that $S \geqslant + D$ (since the difference between b_1 and b_{21} is positive and discernible). The two hypotheses have an equal initial probability, but h_{11} makes our evidence more likely than does h_9: that evidence therefore reciprocates by making h_{11} more probable than h_9, and similar reasoning will show h_{12} to be more probable than h_8, h_{13} than h_7, h_{14} than h_6, and so on. We have seen that this is sufficient to prove p_{i+1} to be more probable than p_i (the probability that b_{i+1} is hotter than b_i is approximately 0.605)[3], so if Jim is rational, although his epistemic preferences amongst p_1, p_2, p_3 etc. will change with the new evidence, there is nothing to suggest that they will cease to be transitive. Sorensen's second example is therefore no more successful than his first: he has given no reason to question the transitivity of epistemic preferability, and hence no reason to reject an interpretation of that notion which conforms to the calculus of probabilities.

---

[2] To find these probabilities mathematically, we again use the Central Limit Theorem to show that h_n yields an approximately normal distribution function for S, with mean $(n-10)D$ and variance $5D^2/3$. h_9 and h_{11} thus give figures of 0.061 and 0.5 respectively for the initial probability that $S \geqslant + D$.

It might perhaps be thought that the above argument itself relies on probabilistic assumptions such as Bayes' Theorem, and thus presupposes a probabilistic (and therefore transitive) conception of epistemic preferability. This objection, however, would be misguided, since the point of the argument is not to establish that epistemic preferability *is* transitive, but only to refute Sorensen's attempt to prove that it is *not*. His 'counterexample' can be accommodated perfectly well within a probabilistic interpretation of epistemic preferability: it therefore fails to overthrow that interpretation.

One final comment. Sorensen hopes to find cases where contradictions arise from the supposition that epistemic preferability is probabilistic. Unless probability theory is actually inconsistent, however, it is difficult to see how he can possibly succeed unless he can show that the relevant basic methods of acquiring information (e.g. by putting hands in buckets of water) themselves yield results which resist probabilistic interpretation. To take Sorensen's first example: as soon as it is agreed that the probability of a research group's success is proportional to its resources, it requires only elementary mathematics to show that the transitivity of epistemic preferability has not been undermined. Similarly with his second example: if the assumptions about initial probabilities given earlier are accepted (and Sorensen has given no reason whatever why they should not be), then we can be sure that probability theory will be able to deal with the apparent contradiction, even if our ignorance of the Central Limit Theorem and of Bayes' Theorem prevents us from performing the necessary calculations ourselves. Thus if Sorensen wishes to show that epistemic preferability is not transitive, he must prove that at least some basic forms of evidence are irreducibly non-probabilistic. How he is to go about doing so is another question!

Lincoln College,
Oxford OX1 3DR

© PETER MILLICAN 1983

---

[3] The complete calculation for $P(d_i > 0)$ turns out like this, ignoring in the intermediate stages values of n which make a negligible difference. Three decimal places are given, but more accurate figures are used in the calculation.

a Initial $P(S \geqslant +D)$ $= 0.349$

b	n:	8	9	10	11	12	13	14	15	16	17
c	Initial $P(h_n)$:	0.120	0.160	0.176	0.160	0.120	0.074	0.037	0.015	0.005	0.001
d	$P(S \geqslant +D)$, given h_n):	0.010	0.061	0.219	0.500	0.781	0.939	0.990	0.999	1	1
e	Final $P(h_n)$ $= c \times d/a$:	0.003	0.028	0.111	0.229	0.269	0.199	0.105	0.042	0.013	0.003
f	$P(d_i > 0$, given $h_n)$ $= n/20$:	0.400	0.450	0.500	0.550	0.600	0.650	0.700	0.750	0.800	0.850
g	Final $P(h_n$ and $d_i > 0) = e \times f$:	0.001	0.013	0.055	0.126	0.161	0.129	0.073	0.032	0.011	0.003

h Final $P(d_i > 0)$
$=$ sum of row g
$= 0.605$

KNOWING AND NON-ACCIDENTAL GUESSING

By TERRY DARTNALL

IN 'Knowing, Believing, and Guessing' (ANALYSIS 42.2, October 1982) Roy Sorensen criticises Colin Radford's argument that knowledge doesn't entail belief ('Knowledge — By Examples', ANALYSIS 27.1, October 1966). Sorensen suggests that Radford's argument rests on principle (A):

> '(A) If one merely guesses that p, then one can only be accidentally right that p.' (p. 212).

Sorensen argues that A is false, and that Radford's argument against the entailment thesis therefore fails. He argues against (A) by arguing for (B):

> (B) One can guess and get it right *non-accidentally*,

i.e. it's possible to guess and get it right time and again without this being luck, accident or fluke.

My sympathies are for the most part with Radford, against Sorensen — so I shall employ what I hope is the disarming tactic of accepting (B) and arguing that, far from *undermining* Radford's thesis, it really *underpins* it. Radford needs non-accidental guessing, and in demonstrating that such a thing is possible. Sorensen does him a service.

I

Radford argues against the entailment thesis (that A knows that p entails that A believes that p) by counterexample. He imagines a situation where Jean, a French-Canadian, believes himself to be ignorant of English history but guesses the answers to a surprising number of English history questions. He then remembers that he once *did* learn some English history.

Radford concludes that Jean must have known the answers all along (even though he did not believe them). He gives two reasons. Firstly, he says, the semantics of 'know' is a matter of ordinary usage, and ordinary speakers would tend to say that Jean *knew*. Secondly, Jean is either guessing or remembering, and is correct 'sufficiently often to preclude the possibility that he might simply be guessing . . . So he *remembers* some history, and hence he knows some history' (p. 5).

Sorensen ignores Radford's first argument and concentrates on his second, that Jean couldn't have guessed right every time by accident, and so couldn't have been guessing at all. Sorensen rejoins that because of the possibility of non-accidental guessing Jean *could* have guessed and got it right time and again, so that the guessing possibility cannot be ruled out. In order to establish the possibility of non-accidental guessing he tells us about Julia. Julia sits for a test consisting of ten problems requiring viscosity calculations to the third decimal place. She uses a method of calculation which she *believes* is only reliable up to the second decimal place, but which is *in fact* reliable up to the third. Consequently, to her astonishment, she gets all her calculations right. Sorensen concludes that this is a case of guessing and getting it right *for a reason* (because the method of calculation is correct to the third decimal place), i.e. is a case of non-accidental guessing.

Well, what *does* happen in Julia's case? She guesses and gets it right so consistently that it is vastly improbable that she is getting it right by chance. Consequently it is vastly probable that there is a *reason* why she keeps getting it right. And there *is* a reason: her method is accurate to three places after all.

Now look at Jean's case. He, too, gets it right so often that it is vastly improbable that he does so by chance. That is the first conclusion we should come to – not Radford's conclusion, *that Jean could not have been guessing*, for Jean quite clearly *is* guessing. Jean says the first thing that comes into his mind whilst not wittingly believing it, and if that ain't guessing, nothing is. No, Jean, like Julia, *guesses* and gets it right so often that it is vastly improbable that he is getting it right by accident, and so vastly probable that there is a reason for his success. Sorensen makes no attempt to explain this success, but I would have thought the answer is fairly clear: Jean guesses and gets it right *because he once learned English history and the memory lingers on*. Julia guesses and gets it right because her method is good to three decimal places. Jean guesses and gets it right because he *remembers* some of what he learned. It's the fact that he *remembers* that explains why his guessing is so successful, that makes it non-accidental in this case.

Thus guessing and remembering are consistent, as Radford needs them to be, for Jean both knows *and* guesses.

II

Why should Radford deny that Jean is guessing? For one thing the form of his argument requires him to say this: (P1) Jean is either guessing or remembering; (P2) Jean is not guessing; therefore (C) Jean is remembering (and so knows). Sorensen correctly observes that Radford's grounds for asserting (P2) are inadequate, because Jean might have been non-accidentally guessing (as, indeed, he was). But Sorensen's observation is unnecessary, since (P2) is false *anyway*: Jean *is* guessing, as we have seen.

I suspect that a source of confusion here is the notion of 'merely' guessing (Sorensen, op. cit., p. 212; Armstrong, 'Does Knowledge

Entail Belief?', *Proceedings of the Aristotelian Society*, October 1969, pp. 25, 27), 'just' guessing (Radford, op. cit., pp. 2, 3), 'simply' guessing (ibid., p. 5), and so on. Sorensen correctly summarises Radford's conclusion by saying that Jean could not have been 'merely guessing' (p. 212), though Radford's actual term is 'simply'. Armstrong says 'It is concluded that his "guess", although it appeared to himself at the time to be a *mere* guess [my emphasis], was in fact something he remembered' (op. cit., p. 25). I think that locutions containing 'merely', 'simply' and 'just' are frequently ambiguous. To say Jean *merely guessed* something might mean either (a) or (b): (a) Jean guessed it, which is a pretty mere thing to do ('I mean, guessing's a bit feeble, isn't it?'); (b) Jean only guessed it – that and no more. Merely guessing in sense (a) is compatible with Jean's *also* having known it; but merely guessing in sense (b) is not. Compare 'James *merely* (sense (a)) stole some string, my dear, and stealing string's the merest thing'. 'Stealing string's the merest thing, but he didn't *merely* (sense (b)) steal the string – he strangled the cat with it.' Radford's original argument states that Jean could not have been *simply* guessing, but 'simply' is equivocal here. It is vastly improbable that Jean was *simply* guessing in the sense of guessing *and nothing else* (sense (b)): but he *was* simply guessing in the sense that guessing's a mere and simple thing to do. However, he was also *remembering*, which is compatible with guessing in this sense (and which explains why he guessed and got it right).

If it is mistakenly assumed that all guessing is 'mere' guessing in sense (b), it follows that guessing p and knowing p are incompatible. This is false, as Jean's case shows: what are incompatible are guessing p and wittingly (consciously) believing p: you can't guess something that you wittingly believe. And, of course,. realizing that guessing and knowing are compatible whereas guessing and believing are not, forces us to abandon the entailment thesis, for if guessing and knowing are compatible then guessing is compatible with anything entailed by knowing, including believing. Since believing is incompatible with guessing, it follows that believing is not entailed by knowing.

III

Where does this leave Radford? His first argument about ordinary usage is still intact. We have not even considered it. His second argument now fails as it stands because (P2), that Jean isn't guessing, is false. But we can easily reconstruct the argument in terms of explanation and plausibility. Jean keeps guessing and getting it right. It is vastly improbable that this is an accident, so what is the most plausible explanation? Since Jean once learned some English history, the most plausible explanation is that he still remembers some, that the memory lingers on. Consequently he knows some, despite his lack of belief.

National University of Singapore,
Kent Ridge, Singapore 0511

© TERRY DARTNALL 1984

THE IMPORTANCE OF BEING COMPLETELY WRONG

By ROY A. SORENSEN

IN 'Knowing, Believing, and Guessing' (ANALYSIS 42.4, October 1982) I offered a defence of the entailment thesis 'Knowledge implies belief' against Colin Radford's influential 'counterexample'. The case involves Jean, a French-Canadian whose reluctant guesses to English history questions are so accurate that we feel compelled to credit him with knowledge despite his lack of belief. After suggesting that the persuasiveness of the example is due to the belief that

> (A) If one merely guesses that p, then one can only be accidentally right that p.

I presented a counterexample to (A) involving another unconfident examinee, Julia. For lack of a better alternative, Julia uses a method of calculation which she believes only assures her a 0.1 probability of a correct answer. However, the method actually assures her of a 1.0 probability of a correct answer. Despite the fact that Julia was merely guessing, her answers are non-accidentally right. I supported my intuition that Julia lacks knowledge with the following sufficient condition for ignorance:

> (B) If one falsely believes that one's method is unreliable, then one does not know.

Since Jean falsely believed that his method was unreliable, (B) also disqualifies Jean from knowing.

In 'Doubting One's Methods' (ANALYSIS 43.3, June 1983), Alan White criticizes my defence. Professor White first objects that (B) is really a variant of the entailment thesis and thus concludes that the defence is question-begging. Since White rejects the entailment thesis, his second objection is that (B) is false.

Is (B) a variant of the entailment thesis in any sense which makes my defence question-begging? I think not because one could deny the entailment thesis without denying (B). For example, some philosophers have denied the entailment thesis on the grounds that the objects of belief are not the objects of knowledge. Others deny it on the grounds that it is inappropriate to assert 'I believe that p' when one knows that p. If my defence of the entailment thesis were question-begging, it would force all opponents of the entailment thesis to reject (B). In fact, my defence is restricted to Radford's objection; other defences are needed against the other objections.

White offers three 'counterexamples' to (B). The first involves someone who has learned how to determine whether certain large numbers are divisible by 4 by checking whether the last two digits are divisible by 4. A mathematician then misleads him into believing that the method is unreliable. White maintains that if our victim should continue to answer questions with this method, he would continue to know the answers (in violation of (B)). I am afraid that all I can say about this case is that my intuition conflicts with White's: it seems clear to me that the misinformed individual does not know the answers.

However, our intuitions are not in as strong a conflict in the remaining cases. These involve sceptics, one sceptical about telescopes and the other sceptical about perception. Philosophers are familiar with the perception sceptic. He will state that perception is an unreliable approach to reality – despite behaving in a way that shows he has perceptual knowledge. As the perception sceptic propounds his thesis, he will do things like reach for his beer, light his pipe, and wave to a friend. So it appears that the sceptic knows, for example, that he has a pipe in his mouth while falsely believing that his method of acquiring this knowledge is unreliable.

However, the same behaviour which reveals perceptual knowledge reveals belief in the reliability of perception. On this basis, one might claim that the sceptic is being insincere, that he is no true sceptic at all; actions speak louder than words.

But we might be dealing with a sophisticated sceptic. The sceptic might concede that he has a normal set of beliefs, including a belief in the reliability of perception. In addition to this normal set, the sceptic claims to have a set of philosophical beliefs which conflict with some of his normal beliefs. The normal beliefs are due to a variety of natural forces which overwhelm reason, giving rise to normal behaviour. Despite the fact that the sceptic's philosophical beliefs are almost entirely restricted to verbal expression (where they do little harm), they are beliefs none the less – not mere talk. Thus the sceptic admits to being inconsistent. He has the natural belief in the reliability of perception in conflict with his philosophical belief in its unreliability. As long as we are willing to protect the sceptic's disbelief in the reliability of perception by maintaining that his conflicting behaviour only shows doxastic inconsistency, we have a counterexample to (B).

One might be tempted to argue that we should not be willing to engage in this protection, but to do so would be needlessly controversial. For (B) can be weakened to avoid White's troublesome cases:

(B′) If one has only incorrect beliefs about the reliability of one's method, then one does not know.

The sophisticated sceptic's claim of leading a doxastic double life is troublesome because he would not be *completely* mistaken about the reliability of his method. He would be right by virtue of his natural belief and wrong by virtue of his philosophical belief. However, Jean and Julia are consistent and are therefore completely wrong about the reliability of their methods. And by virtue of the purity of their error, they are ignorant.

Michigan State University,
East Lansing, Michigan 48823, U.S.A.

© Roy A. Sorensen 1984

TREES FOR A 3-VALUED LOGIC

By Fred Johnson

In [3] Slater uses trees to restrict the classical propositional logic to avoid the paradoxes of material implication and other odd features of the classical logic such as these:

(1) Something (a tautology) follow from nothing

and

(2) Something follows from a contradiction.

The paradoxes of material implication are avoided by restricting the synthetic tree rule[1] that would permit one to write 'p v q' in a path that contains 'p'. (1) and (2) are avoided by requiring that every deduction tree have, respectively, a non-empty and a non-contradictory initial list. Though Slater's project is a good one there are some problems with his discussion. For example, as noted above, he says that (2) is false in his system, but he also says that his system could be presented as a natural deduction system in which there is the Rule of Simplification. With this rule 'p' follows from 'p & ~ p', and thus (2) is not false. Another problem is that he provides no semantic account of validity. I will present a modified tree account of validity and a semantic account of validity with the same extension. We will see that with this account of validity we can avoid the paradoxes and oddities that Slater wished to avoid.

---

[1] Throughout the discussion I will follow the terminology of [1], and I will mime some of the arguments of [1], though only classical logics are developed in [1].

Consider a language in which statements are constructed in the standard way from the letters A_1, A_2, A_3, . . . , parentheses, and the connectives '&' and '~'. To construct trees we will use the following (standard) analytic rules:

$$\begin{array}{c} p \\ \underline{\sim p} \\ x \end{array} \qquad \begin{array}{c} \underline{\sim\sim p} \\ p \end{array} \qquad \begin{array}{c} \underline{p \,\&\, q} \\ p \\ q \end{array} \qquad \begin{array}{c} \underline{\sim (p \,\&\, q)} \\ \sim p \mid \sim q \end{array}$$

Moreover, we will use one (non-standard) synthetic rule:

$$\begin{array}{c} \underline{\quad p \quad} \\ q \mid \sim q \end{array}$$

To use the synthetic rule the initial list has to contain the letters that occur in q. To construct a *deduction tree* we will require that the synthetic rule is not used before all uses of the analytic rules have been exhausted in the tree under construction. Let us call the initial part of the deduction tree that cannot be extended without the use of the synthetic rule the *initial analytic part* (IAP) *of the deduction tree.* Any part of a path P of a deduction tree D that occurs in the IAP of D will be in the *IAP of P of D.* A path P of D will be *open* if 'x' (the 'x' of the analytic rule mentioned first) does not appear in the IAP of P of D; otherwise, it is *closed.* We will say that $P_1, \ldots, P_n$ *syntactically entails* C ($P \vdash C$) if and only if there is a deduction tree with initial list $P_1, \ldots, P_n$ such that (1) there is an open path in the deduction tree, and (2) C occurs in every open path.

To define semantic entailment we will first define valuations as functions which map statements into the set {0, 1, 2}. (0, 1 and 2 can be thought of as false, odd (or meaningless) and true, respectively.) A valuation V meets these conditions: ($V(\sim p) = 2 - V(p)$; $V(p \,\&\, q) = 1$ if either $V(p) = 1$ or $V(q) = 1$, and, otherwise, $V(p \,\&\, q) =$ the minimum of $V(p)$ and $V(q)$. (This definition of V is attributed by Rescher to Bochvar. See p. 29 of [2].) We will say that $P_1, \ldots, P_n$ *semantically entails* C ($P \models C$) if there is a valuation V such that $V(P_1 \,\&\, \ldots \,\&\, P_n) = 2$, and if for any valuation V $V(C) = 2$ if $V(P_1 \,\&\, \ldots \,\&\, P_n) = 2$.

That the notions of syntactic and semantic entailment have the same extension is a corollary of the following four claims.

1. If $P \vdash C$ then there is a valuation V such that $V(P) = 2$. Proof: assume the antecedent. Then there is an open path OP in a deduction tree D with P as its initial list. Let V be a valuation that assigns 2 to those statement letters that are full lines of the IAP of OP and let V assign 0 to all of the other statement letters that occur in the IAP of OP. Any line in the IAP of OP is either part of the initial list or was placed by one of the analytic rules. Since for each

of the analytic rules a 2 below the line guarantees a 2 above the line, it follows that $V(P) = 2$.

2. If $P \vdash C$ then $V(C) = 2$ if $V(P) = 2$. Proof: assume that $V(P) = 2$. If the upper part of an analytic rule is assigned the value 2 then at least one branch below the line is assigned the value 2. The same holds for the synthetic rule. For if $V(P) = 2$ then any statement composed of letters in P must have the value 0 or 2. (For suppose there were a letter in P that had the value 1. Then P would have the value 1. But if each letter in a statement has the value 0 or the value 2 then so does the entire statement.) But then either the 'q' or the '$\sim$q' of the synthetic rule has to have the value 2. So if $V(P) = 2$ then in any deduction tree with P as initial list there is a path in which each full line is assigned the value 2. Such a path must be open. So if we assume that $P \vdash C$ it follows that C is in this path, and thus $V(C) = 2$.

3. If $P \models C$ then there is an open path in any deduction tree that has P as its initial list. Proof: assume the antecedent. Then there is a valuation V such that $V(P) = 2$. By an argument in the preceding paragraph it follows that there is a path in the tree in which every full line is assigned 2 by V. But this path has to be an open path.

4. If $P \models C$ then there is a deduction tree with initial list P and with C in every open path. Proof: first note that if $P \models C$ then any deduction tree with P and $\sim$C as the initial list has no open paths in its IAP. For suppose there were such a path. Then by the argument for the first of the four claims under consideration there would be a valuation V such that $V(P) = 2$ and $V(\sim C) = 2$. But then it would be false that $P \models C$. Keeping this in mind, we use the following recipe to construct a tree:

(a) Construct the IAP of a deduction tree with P as its initial list.
(b) Apply the synthetic rule to every open path, putting 'C' to the left and '$\sim$C' to the right. (Any '$\sim$C' and anything that extends from it will be said to be in a 'right branch'.) Note that there is no difficulty in meeting the qualification on the synthetic rule since if C contained a letter S that did not occur in P then a valuation V that assigns 2 to P but 1 to S would assign 1 to C. Thus, it would be false that $P \models C$.
(c) Apply analytic rules to the 'right branches' until no more applications are possible.

Since each of the 'right branches' is closed (by the argument we are keeping in mind) it follows that C occurs in every open path.

That syntactic and semantic entailment have the same extension, for arguments with one premise, follows directly from statements 1–4. There is no difficulty in generalizing the argument to cover arguments with a greater number of premises.

Let us complete our discussion by looking at some of the puzzles Slater mentioned in [3]. To separate the notion of entailment discussed above from classical entailment, let us call it superentailment. Since if V is any valuation $V(A_1 \,\&\, {\sim}A_1) = 0$ or $V(A_1 \,\&\, {\sim}A_1) = 1$, '$A_1 \,\&\, {\sim}A_1$' does not superentail anything (though it entails everything). Since there is a valuation V such that $V(A_1) = 2$, $V(A_2) = 1$ (and thus $V({\sim}(A_2 \,\&\, {\sim}A_1) = 1)$, it follows that it is false that $A_1 \models {\sim}(A_2 \,\&\, {\sim}A_1)$. Thus, not everything superentails that everything 'materially implies' it (though everything entails that everything materially implies it).

Colorado State University,
Fort Collins, Colorado 80523, U.S.A.

© FRED JOHNSON 1984

REFERENCES

[1] Richard C. Jeffrey, *Formal Logic: Its Scope and Limits*, 2nd ed., New York: McGraw-Hill, 1981.
[2] Nicholas Rescher, *Many-Valued Logic*, New York: McGraw-Hill, 1969.
[3] B. H. Slater, 'Direct Tableaux Proofs', *Analysis*, 41.4 (October 1981), 192–4.

THE EFFECTS OF PRINCIPLES AND OF ACTIONS

By M. C. GEACH

> When a man attempts to combat the principle of utility, it is with reasons drawn, without his being aware of it, from that very principle itself. His arguments, if they prove anything, prove not that the principle is *wrong*, but that, according to the applications he supposes to be made of it, it is *misapplied*.

THIS passage is taken from Bentham's *Principles of Morals and Legislation*, where he defines the principle of utility as 'that principle which approves, or disapproves, of every action whatsoever according to the tendency which it appears to have to augment or diminish the happiness of the party whose interest is in question'.

We should notice that this principle is one which judges actions rather than principles. (I assume, however, that Bentham thought of it as a principle of conduct, to be employed as much in the judgement of *proposed* actions as in the judgement of actions already performed.)

In a footnote, Bentham mentions a man who calls the principle of utility 'dangerous'. Such a man, he thinks, must be arguing with reasons drawn from the principle of utility itself. 'Dangerous', Bentham thinks, must mean 'tending to reduce happiness', and so the principle can only be attacked as dangerous when the 'party whose interest is in question' is different from the party which the principle is said to endanger.

Evidently, Bentham supposes that the principle of utility, when not misapplied, must tend rather to increase happiness than to diminish it. Lying behind this supposition is the plausible assumption that people who go by a principle which judges actions for their tendency to produce some kind of effect must tend to produce more of that kind of effect than do people who prefer some other principle. I call this assumption 'Bentham's Assumption', not because he explicitly states it, but because it seems the only plausible ground for his criticism of those who object to the principle of utility as one which tends to diminish happiness.

However, Bentham's Assumption is disproved by the case of the prisoners' dilemma. This famous example goes as follows:

Two prisoners are being asked to inform, each on the other. Each is told that if he informs and the other does not, then the other will be given ten years but he himself will be set free. The opposite would happen if he kept mum and the other informed on him. If each informs on the other, they will each get five years; but if neither informs on the other, they will each get two, for the police will see that they are convicted on another charge. Neither prisoner's decision will affect the decision made by the other.

If the prisoners both go by the principle 'Act in the way which tends to minimize your sentence', they will betray one another. For each thinks 'In the event that the other betrays me, I will be better off betraying him. For then I will only get five years, while silence would get me ten. But in the event that the other does not betray me, I will be better off betraying him, for then I will go scot-free, while silence would get me two years. Since my decision will not alter his, there is no doubt what I should do.'

One might think that the prisoners cannot each be correctly applying the principle that he should act in a way tending to increase his liberty: for the pair who prefer the principle 'Do not betray' only suffer two years apiece, which is a shorter sentence than would be suffered by the pair each of whom act on the principle that he should act in the way tending to minimize his sentence.

However, they have not misapplied this principle, for each of a pair of betrayers can congratulate himself on having escaped a ten-year sentence, and can despise each of a pair of loyal men, for each, given the silence of the other, could have got away scot-free. Thus people who go by a principle which judges actions for their tendency to produce a kind of effect do not necessarily tend to produce more of that kind of effect than do people who prefer some other principle. For each of a pair of traitors has less freedom than does each of the pair who are loyal.

This case might be thought to have its peculiar properties because it is a case of selfishness; the effects aimed at by the traitors are different and incompatible effects, and it is not surprising that, as a pair, the traitors should defeat one another.

However, a similar sort of case can arise even when people agree about the desired effect, and are not selfish. We find such a case in the well-known puzzle of the voter, which I embellish as follows: in a referendum, there is a question of whether some measure is to be passed which appears to have a tendency to augment the happiness of mankind at large. The act-utilitarians all want the measure passed, for they desire the happiness of mankind at large. However, the act-utilitarians go by the principle that actions are to be judged by their tendency to augment or diminish the happiness of mankind at large. They therefore do not vote, for people consistently acting on this principle will hardly ever vote. A single vote is practically certain to make no difference to the referendum, and so the act of voting appears to have practically no tendency to augment the happiness of mankind at large, unless for some accidental reason, or because the voter likes voting. Thus few consistent act-utilitarians vote and, in consequence, the referendum goes against them.

This example also shows that Bentham's Assumption is not true, and so it might be that the principle of utility (even when not misapplied) is not only dangerous but actually destructive of human happiness. Thus Bentham's criticism of those who call the principle of utility 'dangerous' fails to make its point. But the wrongness of what he says, once seen, may show us the way to a study of ethics which will take better note of the intricacy and variety of the things which we call 'ends', and of the ways in which these specify our actions.

3 Richmond Road,
Cambridge

© M. C. Geach 1984

RATIO edited by MARTIN HOLLIS

Volume XXV (1983) **December**

ROGER SCRUTON	Understanding Music
HEINZ-DIETER HECKMANN	What a *Res Cogitans* is – and Why I am one
LAURENCE GOLDSTEIN	Wittgenstein and the Logico-Semantical Paradoxes
GEOFFREY BROWN	Language, Rules and Authority
E. M. ZEMACH	The Possibility of a Materialistic Language
ROBERT BERNASCONI	Review of Rudiger Bubner, ***Modern German Philosophy***
DAVID WOOD	Review of Vincent Descombes, ***Modern French Philosophy***
NICHOLAS EVERITT	Review of Roderick Chisholm, ***The Foundation of Knowing***

Subscriptions
Institutions. £22.95 (UK) £28 20 (overseas) $59.95 (USA) $74 00 (Canada)
Individuals £15.50 (UK) £19.00 (overseas) $42 00 (USA) $52.00 (Canada)

Orders to. Journals Department, Basil Blackwell Publisher Ltd , 108 Cowley Road, Oxford OX4 1JF, England.

The German edition of *Ratio* is published by FELIX MEINER VERLAG, Richardstrasse 47, 2000 Hamburg 76, West Germany

Published by BASIL BLACKWELL, Oxford

INQUIRY *Editor: Alastair Hannay*

Vol. 26 (1983), No. 2

Nietzsche and the Origin of the Idea of Modernism	Robert B. Pippin
Values Gone Wild	Holmes Rolston, III
Waking-up: A Neglected Model for the Afterlife	Stephen R.L. Clark
Mathematical Modeling of Election Predictions: Final Reply to Professor Aubert	Herbert A. Simon
Nature May Be of No Value: A Reply to Elliot	Don Mannison
Plato on the Sciences (J.P. Anton (Ed.), *Science and the Sciences in Plato)*	Georgios Anagnostopoulos

Vol. 26 (1983), No. 3

Noumenalism and Einstein's Argument for the Existence of God	Lewis S. Feuer
Marx and Cohen on Exploitation and the Labor Theory of Value	Nancy Holmstrom
More on Exploitation and the Labour Theory of Value	G.A. Cohen
The Epistemological Status of a Naturalized Epistemology	Ron Amundson
Interpreting Frege: A Reply to Michael Dummett	Gregory Currie
The Value of Wild Nature	Robert Elliot
Recent Work on Frege (G. Currie, *Frege: An Introduction to His Philosophy;* Hans Sluga, *Gottlob Frege;* M. Dummett, *The Interpretation of Frege's Philosophy)*	Crispin Wright

Subscription price:
Institutions: N.kr. 185,— $32.00 — Individuals N.kr. 110,— $19.00.

Published quarterly by Universitetsforlaget, P.O. Box 2959 Tøyen, Oslo 6, Norway — P.O. Box 258, Irvington-on-Hudson, New York 10533, USA.

Philosophical Relativity

PETER UNGER

'This is an intelligent and highly original critique, clearly and even gracefully written, with a refreshing absence of pedantry.' P.F. Strawson

'First rate philosophy, philosophy as it ought to be done.' Gilbert Harman

It is still generally believed that there are definitive objective answers to the traditional problems of philosophy. In this book Peter Unger questions this widespread belief.

260 pages, c.£17.50 (0 631 13428 X)

Hegel: An Introduction

RAYMOND PLANT

Hegel could well be the most misunderstood of all philosophers. His complex metaphysics, confusing terminology and his almost impenetrable style have discouraged generations of readers. And yet Hegel is undoubtedly one of the most profound and important philosophers of modern times. In this lucid introduction to Hegel's thought, Raymond Plant uses his political and social writings as a key to the rest of his work.

254 pages, hardback £17.50 (0 631 13333 X)
paperback £6.50 (0 631 13334 8)

Rights, Killing and Suffering

Moral Vegetarianism and Applied Ethics

RAYMOND FREY

Is it morally wrong to eat meat? In recent years much has been written about animal welfare and rights, and more and more people have come to think that vegetarianism is the only justifiable approach. Their views reflect one or more of three basic arguments: that meat-eating violates moral rights, that it is wrong to kill, and that it is wrong to inflict pain and suffering. In this provocative book, going against the main tide of current thinking, Raymond Frey examines and finds wanting all three arguments.

268 pages, £17.50 (0 631 12684 8)

The Roots of Knowledge

NATHAN STEMMER

How do we acquire our initial beliefs about the nature of the world? How reliable are those beliefs? How do we pick up the rudiments of language? How can these crude beliefs and elementary linguistic tools be reliable guides to the external world? This stimulating book advances a novel theory to explain our acquisition of first subjective beliefs and their function as a basis for the later acquisition of objective knowledge.

208 pages, £17.50 (0 631 13328 8)

Basil Blackwell

New in paperback

Wittgenstein on Rules and Private Language

SAUL A. KRIPKE

'Saul Kripke has thought uncommonly hard about the central argument of Wittgenstein's **Philosophical Investigations** and produces an uncommonly clear and vivid account of that argument . . . clearly and compellingly presented . . . an exemplary piece of exposition.' P.F. Strawson, **Times Literary Supplement**

160 pages, hardback £9.50 (0 631 13077 2)
paperback £5.50 (0 631 13521 9)

Wittgenstein and his Times

Edited by BRIAN McGUINNESS

Wittgenstein and his Times contains five essays by well-known philosophers – Anthony Kenny, Brian McGuinness, J.C. Nyiri, Rush Rhees, G.H. von Wright – who relate his thinking and world view to that of certain influential contemporaries.

128 pages, hardback £12.00 (0 631 11161 1)
paperback £4.95 (0 631 12925 1)

Knowledge and the Flow of Information

FRED I. DRETSKE

'I recommend his book . . . presents the relevant figures of information theory clearly and is philosophically sophisticated.' **Philosophical Books**

'The book sparkles with developed insights. Its flow of information is not flawless, but it is certainly forceful and fruitful.' **Times Literary Supplement**

288 pages, hardback £16.50 (0 631 12765 8)
paperback £7.50 (0 631 13305 4)

Illegitimacy

A Philosophical Examination

JENNY TEICHMAN

'Undergraduates wishing to study some specific aspect of illegitimacy can make an excellent start with this volume.' **Choice**

Jenny Teichman's book is a new departure – a philosophical analysis of the very notion of illegitimacy, tracing its meaning and importance in the practices of different societies and in the tenets of different creeds.

208 pages, hardback £16.50 (0 631 12087 7)
paperback £6.50 (0 631 13588 X)

Basil Blackwell

The Philosophical Quarterly

Vol.34 No.134 JANUARY 1984
© The Philosophical Quarterly 1984

Edited in the University of St. Andrews by an Editorial Board
Chairman of the Board of Editors: LESLIE STEVENSON
Executive Editor: GORDON GRAHAM *Reviews Editor*: ANTHONY ELLIS

Emotion and Imagination	*John Casey*
Vico's Theory of Knowledge	*Perez Zagorin*
Dummett on Temporal Operators	*Jeremy Butterfield*
Nozick on Scepticism	*Graeme Forbes*
DISCUSSIONS	
Conjectural Realism	*Richard Jennings*
Brown in Defence of Scientific Realism	*Roland Puccetti*
Reply to Puccetti	*J. R. Brown*
Mill and the Utility of Liberty	*Mark Strasser*
BOOK REVIEWS	

Subscriptions for 1984

Annual Subscription	UK	Overseas	USA	Canada
Institutions	£22.50	£27.50	$50.00	$62.00
Individuals	£10.25	£12.50	$25.00	$30.00

Orders, remittances and enquiries should be addressed to: Journals Department,
Basil Blackwell Publisher Ltd., 108 Cowley Road, Oxford OX4 1JF, England.

MIND A QUARTERLY REVIEW OF PHILOSOPHY
edited by D. W. Hamlyn
with the assistance of S D Guttenplan

Vol XCIII No 369 January 1984

Articles

Weakness of Will: F. JACKSON
Fanciful Arguments for Realism: A. H. GOLDMAN
A New Incompatibilism: N. M. L. NATHAN
Doxastic Incontinence: J. HEIL
Subjective Justification: J. L. KVANVIG

Discussions

Taylor on Meaning-Theories and Theories of Meaning: M. DAVIES
Strawson on the Antinomy: L. BALDACCHINO
Is the Prisoner's Dilemma an Insoluble Problem?: D. GORDON
Relevant Interest and the Prisoner's Dilemma: J. P. PORTER
How Incomplete Is Prescriptivism?: H. J. GENSLER
Epistemic Indolence: A Reply to Schmitt: R. FOLEY and R. FUMERTON

Critical Notice

A. MacIntyre, *After Virtue*: D. Z. PHILLIPS

Book Reviews **Books Received** **Notices**

Orders to: Journals Department,
Basil Blackwell Publisher Ltd.,
108 Cowley Road,
Oxford OX4 1JF, England

Annual subscription:
Institutions: £14.50 or $32.50 (Canada $37.75)
Individuals: £10.00 or $22.50 (Canada $26.25)
Single issues: £4.00 or $9.00 (Canada $10.50)

Published for the MIND ASSOCIATION by BASIL BLACKWELL PUBLISHER

NOTES

The ANALYSIS Committee consists of: Chairman, P. T. Geach; Secretary, T. R. Baldwin; J. H. Benson, Margaret A. Boden, A. E. Fisher, Andrew Harrison, R. H. Holland, J. H. McDowell, D. H. Mellor, R. G. Swinburne, A. R. White, C. J. F. Williams, Peter Winch. This committee is responsible for appointing and advising the Editor and for the general policy of the paper.

SUBSCRIPTIONS. ANALYSIS is published by Basil Blackwell Publisher Ltd, 108 Cowley Road, Oxford OX4 1JF, England. Each volume comprises four numbers, each of at least 48 pages, appearing in January, March, June and October. Orders should be sent to the publisher. Subscription prices are as follows:

	U.K.	Overseas	U.S.A.	Canada
Institutions	£11.00	£13.50	$33.00	$39.50
Individuals	£7.00	£8.60	$21.00	$25.00

U.S. MAILING. Second class postage paid at New York. Postmaster: Send address corrections to: Analysis: c/o Expediters of the Printed Word Ltd, 527 Madison Avenue, Suite 1217, New York, NY 10022 (U.S. Mailing Agent).

CONTRIBUTIONS. Articles submitted for publication should be addressed to Mr Christopher Kirwan, Exeter College, Oxford OX1 3DP.

Contributors are asked to note the following requirements. Articles should normally not exceed 3,000 words in length, including references and footnotes. They should be type written in double spacing on one side of the paper only; photographic copies are acceptable. Footnotes should be kept to a minimum and wherever possible avoided altogether. Inset quotations and formulae should be well spaced from the text preceding and following them. The spacing of symbols within formulae should be carefully attended to. Single quotation marks should be used, except for purposes of internal quotation and "scare" quotes. Discussion papers should be sent in as soon as possible after the appearance of the article to which they refer.

Typescripts of rejected articles will be returned only if self-addressed envelopes with correct postage are provided (in the case of overseas contributors, international reply coupons sufficient for airmail).

Proofs of accepted articles will be sent to authors for correction, together with information about free copies. Typescripts of accepted articles will be retained by the Editor on the assumption that authors have kept their own copies.

Apart from correction of printers' errors, only very minor changes in proof will normally be accepted by the Editor, to avoid delay and expense in production.

29 NOV 1985

Typeset by Unicus Graphics Ltd, Horsham. Printed and bound in Great Britain by Whitstable Litho Ltd, Whitstable

Vol. 44 No. 2 (New Series No. 202) March 198

ANALYSIS

Edited by

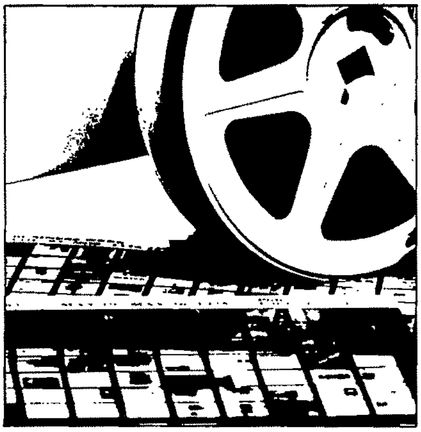
CHRISTOPHER KIRWAN

CONTENTS

On the nature of INUS conditionality	THEODORE C. DENISE
Anscombe on coming into existence and causation	DAVID GORDON
Theseus' clothes-pin	HUGH S. CHANDLER
Putnam's brains	JANE MCINTYRE
An argument in favour of non-classical logic for quantum theory	NICHOLAS HALLAM
Why the distributive law is sometimes false	P. F. GIBBINS
A puzzle about logical equivalence	STOJAN T. DOBROSAVLJEVIĆ
Wright's reply to Benacerraf	MARIE MCGINN
Why physics can't be nominalized	ALAN MCMICHAEL
McGinn on experience of primary and secondary qualities	DAVID MCNAUGHTON
A note on Smith on attempts and internal events	MAURICE RICKARD
Anomalous monism and epiphenomenalism: a reply to Honderich	PETER SMITH
Smith and the champion of mauve	TED HONDERICH
Hempel revisited	A. J. DALE
More on empirical significance	M. L. POKRIEFKA
The tiger and his stripes	WILLIAM LYONS
Green on dictators and democracies	DAVID GORDON

BASIL BLACKWELL PUBLISHER · OXFORD

ISSN 0003-2638



This Publication
is available in Microform.

University
Microfilms
International

Please send additional information
for ______________________
(name of publication)
Name______________________
Institution______________________
Street______________________
City______________________
State______________ Zip______________

300 North Zeeb Road
Dept. P.R.
Ann Arbor, Mi. 48106

Basil Blackwell

Personal Identity

SYDNEY SHOEMAKER and RICHARD SWINBURNE

What does it mean to say that this person at this time is 'the same' as that person at an earlier time? If the brain is damaged or the memory lost, how far does a person's identity continue? In this book two eminent philosophers develop very different approaches to the problem: Sydney Shoemaker presents a materialist and Richard Swinburne a dualist account.
Great Debates in Philosophy
160 pages, hardback c.**£17.50** (0 631 13208 2)
paperback c.**£5.95** (0 631 13432 8)

Frege: Logical Excavations

G.P. BAKER and P.M.S. HACKER

The authors challenge the current interpretation of Frege. By examining the mathematical origins of his thought, they show that his concern was to apply sophisticated ideas from function theory to the logicians' traditional activity of analysing inferences, judgements and concepts. They expose and critically examine his ideas from this viewpoint.
424 pages, **£29.50** (0 631 13169 8)

Philosophical Relativity

PETER UNGER

'This is an intelligent and highly original critique, clearly and even gracefully written, with a refreshing absence of pedantry.'
P.F. Strawson

'First rate philosophy, philosophy as it ought to be done.'
Gilbert Harman

It is still generally believed that there are definite objective answers to the traditional problems of philosophy. In this book Peter Unger questions this widespread belief.
260 pages, **£17.50** (0 631 13428 X)

Liberalism and the Origins of European Social Theory

STEVEN SEIDMAN

A striking reinterpretation of the development of European social theory. Steven Seidman departs from established historical models and returns to the original works of those figures who created modern social theory. He offers a revised view of the Enlightenment and argues for its primacy in the history of social theory. The book makes a significant contribution to intellectual history.
432 pages, **£25.00** (0 631 13452 2)

Basil Blackwell

Consciousness and Causality

A Debate on the Nature of Mind

D.M. ARMSTRONG and NORMAN MALCOLM

How are the objects of consciousness related to consciousness? Can consciousness be defined as awareness of inner mental states? Can one perceive consciousness in oneself? In other persons? Can states of be identified with brain states? David Armstrong and Norman Malcolm here present opposing views on these and related questions in the philosophy of mind.

Great Debates in Philosophy

240 pages, hardback c.**£17.50** (0 631 13212 0)

paperback c.**£5.95** (0 631 13433 6)

Morality and Conflict

STUART HAMPSHIRE

Stuart Hampshire has had an influence far beyond the traditional bounds of academic philosophy. His lucid prose and compelling arguments have made his works essential reading for students and scholars alike, while his concern for practical issues in moral and political philosophy has attracted a wide following among the general public. In these essays he explores the complex relationship between morality and conflict.

136 pages, **£15.00** (0 631 13336 4)

Freud and the Mind

İLHAM DILMAN

This is the first full-scale philosophical treatment of Freud's conception of the mind and of the limits of the individual's autonomy. The author explains the concept of the unconscious, the prominence Freud gave to phantasy and the emotion in human life, and the way he conceived of the relation between reason and the emotions.

224 pages, **£15.00** (0 631 13529 4)

Freud and Human Nature

İLHAM DILMAN

Interest in the writings of Freud seems never to decline, but it is rare that an important book appears which discusses Freud's belief that human nature exists independently from the society in which an individual develops. This book provides this: a thorough critical examination of the various issues raised by Freud's ideas of sexual and moral development.

216 pages, **£15.00** (0 631 13373 9)

Basil Blackwell

Beyond Objectivism and Relativism

RICHARD J. BERNSTEIN

Philosophy is again questioning its goals and assumptions, a malaise signified by the growing opposition between objectivism and relativism, between rationality and irrationality, between realism and anti-realism. Contemporary thinking is becoming increasingly polarized and attempts to escape from this entire framework of analysis have all too frequently been assimilated to these standard oppositions. This major new book represents a courageous attempt to break out of this mould.
302 pages, hardback **£19.50** (0 631 13429 8)
paperback **£7.50** (0 631 13437 9)

Philosophy and Sociology of Science: An Introduction

STEWART RICHARDS

Modern science is inseparable from the society that supports it. Yet now more than ever before the scientist is encouraged to specialize and narrow his field of vision. Stewart Richards argues that it is vital to reverse this trend, and that all students of science should look at the context, roots and structure of their disciplines.
220 pages, hardback **£17.50** (0 631 13411 5)
paperback **£5.95** (0 631 13414 X)

Thinking Matter

JOHN W. YOLTON

This examination from Cudworth to Priestley, of a growing debate over the question 'can matter think', will be essential reading for anyone interested in the origins of modern philosophy and the philosophy of science.
320 pages, c.**£17.50** (0 631 13335 6)

Perceptual Acquaintance

JOHN W. YOLTON

This highly original book challenges the accepted wisdom, derived from seventeenth- and eighteenth-century writers such as Locke and Hume, about epistemology. Philosophers then were deeply concerned with sense perception and its relationship to physical objects.
350 pages, c.**£17.50** (0 631 13431 X)

Basil Blackwell

New in paperback

Wittgenstein

G.H. von WRIGHT

'Will be of great value to Wittgenstein scholarship.' *Times Higher Education Supplement*

The author has collected eight studies of Wittgenstein's life and writings, to provide source material that will be of permanent value to anyone studying this major twentieth-century philosopher.

222 pages, hardback **£14.00** (0 631 13099 3)
paperback **£5.95** (0 631 13598 7)

Wittgenstein on Rules and Private Languages

SAUL A. KRIPKE

'Saul Kripke has thought uncommonly hard about the central argument of Wittgenstein's *Philosophical Investigations* and produced an uncommonly clear and vivid account of that argument . . . clearly and compellingly presented . . . an exemplary piece of exposition.' P.F. Strawson, *Times Literary Supplement*

160 pages, hardback **£9.50** (0 631 13077 2)
paperback **£5.50** (0 631 13521 9)

Wittgenstein and the Vienna Circle

Conversations Recorded by **Friedrich Waismann**
Edited by BRIAN McGUINNESS

'This record, splendidly edited by Brian McGuinness . . . is now presented in English translation. One should be grateful to Blackwell for publishing the English translation of Waismann's notes.' Norman Malcolm, *London Review of Books*

'An important and necessary contribution on Wittgenstein.' *Choice*

266 pages, paperback **£4.95** (0 631 13469 7)

Theology and Ethics

JAMES M. GUSTAFSON

'Professor Gustafson is a highly distinguished scholar, a widely cultured, up-to-date thinker, and a man eager to accept the sound conclusions of scholarship in Biblical studies as in every other field.' *Church Times*

360 pages, hardback **£17.50** (0 631 12945 6)
paperback **£7.50** (0 631 13597 9)

COMPETITION ANNOUNCEMENT

The Metaphilosophy Prize

Metaphilosophy will award a prize of 500 American dollars to the author who submits the best essay in Computer Ethics between January 1st, 1984 and December 31st, 1984.

The prize-winning essay will be published as the lead article in the April 1985 issue of *Metaphilosophy*, which will be devoted entirely to Computer Ethics. Other high quality essays in Computer Ethics will be accepted for publication in the same issue.

A panel of experts in Computer Ethics will select the winners. To enter the competition, send four copies of your essay to:

Terrell Ward Bynum,
Editor, Metaphilosophy
Metaphilosophy Foundation.
Box 32, Hyde Park,
NY 12538-0032
USA

Readers unfamiliar with the field of Computer Ethics should consult the January 1984 issue of *Metaphilosophy*. Those unfamiliar with specifications for manuscript preparation should consult any recent issue.

ANALYSIS 44.2 MARCH 1984

ON THE NATURE OF *INUS* CONDITIONALITY

By THEODORE C. DENISE

IN this brief paper, I propose and support an amendment to the established definition of an *insufficient* but *necessary* part of a condition that is itself *unnecessary* but *sufficient* for a result (an INUS condition).

The success of analysing causal statements – most notably those of ordinary discourse – in terms of INUS conditions is well known. Let it be, for example, that I summon a mechanic because my antique car's engine will not start and that he subsequently reports that externally supplied electrical energy *caused* it to start (he used battery jumper cables). Context considered, what does the causal statement tell us? Chiefly this: the externally supplied electrical energy was an INUS condition of the engine's starting. It was a condition in a certain cluster of actual conditions (including also controlled gasoline flow, dry spark plugs, etc.) that together were sufficient for the desired result. Neither it by itself nor any of the other conditions by themselves would have been sufficient for that result. Furthermore, there are other such clusters of conditions (one includes cranking) that would result in the engine's starting. The externally supplied electrical energy was singled out as *the* cause in the successful cluster since, so to say, when it was introduced an otherwise insufficient actual cluster became the sufficient actual cluster.

I shall draw on J. L. Mackie's two classic articles 'Causes and Conditions' (*American Philosophical Quarterly* 2, 1965, pp. 245–64) and 'Mill's Methods of Induction' (*The Encyclopedia of Philosophy*, ed. P. Edwards, 1967, vol. 5, pp. 324–32) and on Frank Jackson's belated review of the former (*Journal of Symbolic Logic* 47, 1982, pp. 470–3). The respective articles have served to standardize discussions of causes as INUS conditions and the eliminative methods of induction. The review demonstrates that Mackie's definition – the established definition – of an INUS condition is too broad.

I

'*A is an INUS condition of a result P if and only if, for some X and for some Y, (AX or Y) is a necessary and sufficient condition of P, but A is not a sufficient condition of P and X is not a sufficient condition of P*' (Mackie, 'Causes and Conditions' p. 246).

Putting us on Boolean alert, as it were, Jackson shows that this, the established definition, suffers intuitively irrelevant conditions in its denotation. Let 'A' be, say, 'My office is in the Hall of Languages', 'P' be, say 'No two snow flakes have identical patterns', and 'S_1' and 'S_2' be marks for distinct sufficient conditions of P such that (S_1 or S_2) is necessary and sufficient for P. Now let $X =$ (S_1 or $\bar{A}$) and $Y =$ (S_2 or $\bar{A}S_1$) and note the following: 'AX is a sufficient condition for P because if it obtains, then S_1 does [if A(S_1 or $\bar{A}$), i.e. (AS_1 or $A\bar{A}$), i.e. AS_1, then S_1]. But A is not a sufficient condition for P, and neither is X. Moreover, (AX or Y) is a necessary and sufficient condition for P, since it obtains if and only if (S_1 or S_2) does [(AS_1 or (S_2 or $\bar{A}S_1$)), i.e. (S_1 or S_2) therefore (S_1 or S_2) $\equiv P$]; and so A satisfies Mackie's definition' (Jackson, op. cit. p. 472).

Can the definition be narrowed without jettisoning its existing structure or without imposing question-begging restrictions on a field F of relevant factors about which, as it happens, we are concerned? Fortunately, it can be. We have merely to append the restriction that *A is a necessary condition of PX* to Mackie's definition. That is, we have merely to supplant the established definition by the following: *A is an INUS condition of a result P if and only if, for some X and for some Y, (AX or Y) is a necessary and sufficient condition of P and A is a necessary condition of PX, but A is not a sufficient condition of P and X is not a sufficient condition of P.*

It is obvious that the amended definition can only narrow the denotation of the established one; and it is demonstrable that it does so. Jackson's move is rendered innocuous. With the foregoing assignments still in place, consider: A is not a necessary condition of PX, i.e. PX can obtain even though A fails to, since, if $P\bar{A}$ obtains, then, although PX, i.e. P(S_1 or $\bar{A}$), obtains, A does not. Thus, A does *not* satisfy the amended definition. We have saved the individuality of snow flakes from INUS dependence on my office location!

II

The need to supplant the established definition by the amended one is greater than just the need to remove a benign anomaly. The dragon of irrelevant INUS conditions has easy access to the garden. To see that this is so, let us regard Mackie's *in abstracto* illustration of one of his novel eliminative methods of induction as if it is illustrative of the INUS concept. (The temporal proximity of his two classical articles suggests that Mackie regarded it in this way.) The *method* is this: 'If there are n possible causes in all, and we observe 2^n instances (positive or negative) which cover all possible combinations of causes and their negations, then the disjunctions of all the conjunctions found in the positive instances is both necessary and sufficient for P in F' (Mackie, 'Mill's Methods of Induction' p. 328). And, with '*a*' for 'absent' and '*p*' for 'present' and a numbered 'S' marking each row where P is present, the illustration is this:

	A	B	C	P	
(1)	*p*	*p*	*p*	*a*	
(2)	*p*	*p*	*a*	*p*	S_1
(3)	*p*	*a*	*p*	*p*	S_2
(4)	*p*	*a*	*a*	*a*	
(5)	*a*	*p*	*p*	*a*	
(6)	*a*	*p*	*a*	*p*	S_3
(7)	*a*	*a*	*p*	*a*	
(8)	*a*	*a*	*a*	*a*	

We note that (i) all and only the complex conditions in the S-rows, viz., $AB\bar{C}$, $A\bar{B}C$ and $\bar{A}B\bar{C}$, are sufficient conditions of P in F and that (ii) each positive or negative component or simple condition ('cause') in each S-row is an insufficient but necessary part of the complex condition (conjunction of simple conditions) in which it occurs. Are facts (i) and (ii) enough to show that each component condition of $AB\bar{C}$, $A\bar{B}C$ and $\bar{A}B\bar{C}$ is, with respect to its complex condition, an INUS condition of P in F? Under the established definition, yes; under the amended definition, no.[1] The latter requires also that (iii) each of these component conditions should be a necessary condition of P conjoined with its associated value of X, with, that is, the remainder of the complex condition in which it occurs. But (iii) is falsified in the illustration by the cases of A in $AB\bar{C}$ and $\bar{A}$ in $\bar{A}B\bar{C}$. Our understanding of what it is to be a necessary condition makes it evident that, if both of these, then neither A nor $\bar{A}$ can be a necessary condition of $PB\bar{C}$ or any of its components.

The logical difference here cannot be ignored. Under the established definition, each conjunct in the disjunctions of the statement $(AB\bar{C}$ or $A\bar{B}C$ or $\bar{A}B\bar{C}) \equiv P$ is, with respect to the conjunction in which it occurs, an INUS condition of P in F. Under the amended definition, only each conjunct in the disjunctions of that statement reduced to $(B\bar{C}$ or $A\bar{B}C) \equiv P$ is, with respect to the conjunction in which it occurs, an INUS condition of P in F.

Both definitions guard against the circumstance of there being an opposed pair of component conditions $(A, \bar{A})$ in a complex condition sufficient for P in F. That is, both guard against contradiction. But only the amended definition guards against the circumstance of there being an opposed pair of component conditions $(A, \bar{A})$ such that, even though one of them is in one of a nominally distinct pair of complex conditions sufficient for P in F and the other is in the other, the distinctness of this pair of complex conditions affording the opposed pair of component conditions separation resides solely in its containing them. That is, only the amended definition guards fully against irrelevance.

---

[1] The Jacksonian assignments for A in row (2) are: $X = (S_1$ or $\bar{A})$, i.e. $(AB\bar{C}$ or $\bar{A}BC$ or $\bar{A}B\bar{C}$ or $\bar{A}\bar{B}C$ or $\bar{A}\bar{B}\bar{C})$; $Y = (S_2$ or S_3 or $\bar{A}S_1)$, i.e. $(A\bar{B}C$ or $\bar{A}B\bar{C}$ or $\bar{A}AB\bar{C})$, i.e. $(A\bar{B}C$ or $\bar{A}B\bar{C})$. *Mutatis mutandis* for the other pertinent component conditions of the complex conditions sufficient for P in F.

Syracuse University,
541 Hall of Languages,
Syracuse, NY 13210, U.S.A.

© THEODORE C. DENISE 1984

ANSCOMBE ON COMING INTO EXISTENCE AND CAUSATION

By DAVID GORDON

ELIZABETH ANSCOMBE has claimed (following Hobbes) that there is a difficulty in supposing that something has come into existence at a particular time and place without supposing that it has a cause.[1] (She thinks, further, that except for the case of the beginning of the world, everything that comes into existence does so at a particular time and place.)

Her argument is the following. If it is supposed that something can come into existence without a cause, then one needs to have reason to say that the thing has 'come into existence at this time and place and not at any other' (p. 160). At first, one might think this poses no problem: can't one, after all, just see that at a certain time t and place p there was nothing present, while at t + 1 there *is* something there?

Professor Anscombe of course recognizes this answer but holds that it fails to resolve the difficulty. 'But we must notice that there is a difference between coming to be at a certain place – that is, its coming about that a thing is at a certain place – and a thing's coming into existence at that place. So the question arises: granted that it has come about that this thing is at this place (where it was not before) why is this a beginning of its existence? Might it not have existed previously elsewhere and arrived here now?' (p. 161).

It won't do, in reply, to say that if the object had come from some other place, we would have seen it arrive. 'I have got to be assured that such objects never behave as we are told α particles do and that this one did not travel from anywhere else in some other form – say as a gas – and assume the form it has here on arrival. The task is too much for me; and for all I have a right to judge in the matter, I am forced to conclude that this object may have come into existence in any place and at almost any time you care to mention. It seems that there is no experience which itself *positively* indicates that I have to do with a beginning of existence here, except indeed that the object is here now and was not here before. But, that being obviously insufficient, I have got to exclude other explanations of its arrival here' (p. 161).

---

[1] G. E. M. Anscombe, 'Times, Beginnings and Causes' in her *Metaphysics and the Philosophy of Mind* (Minneapolis, University of Minnesota Press, 1981), pp. 148–62. All subsequent references to this article will be by page number in the text.

Professor Anscombe suggests that the way we know that objects have begun to exist at a particular time and place rather than arriving from elsewhere is by knowing their causes: 'So in a hundred cases we know we can observe beginnings of new items because we know how they were produced and out of what. It is preposterous to claim that no positive actual or possible experience reveals beginnings of existence. We know the times and places of their beginnings without cavil because we understand their origins' (p. 162).[2]

I do not think this argument succeeds. Exactly the same difficulty as Professor Anscombe raises for knowing whether an object that appears, uncaused, at a particular time and place has really begun there also arises for objects whose causal origins we claim to know. To take one of her examples: suppose we say we know that a pudding begins to exist at a particular time and place because 'we know its origin, for it was made in the kitchen out of its ingredients by your sister, you watched her do it' (p. 162).

But *do* we really know this? How do we know that it isn't the case that the supposed ingredients simply disappear and that the pudding arrives fully complete from somewhere else? On this supposition, we aren't denying that mixing the ingredients causes the pudding to appear: we are denying that the pudding has been made out of the ingredients. If Professor Anscombe is willing to countenance undetectable arrivals out of nowhere as a reason against saying that *uncaused* objects can't be known to begin to exist at a particular time and place, why does the presence of a causal process change anything in the objection?

One might attempt to counter my argument by claiming that I have assumed the truth of a particular view of causation. If one thinks that a cause *necessitates* its effect, isn't the force of my example destroyed? Doesn't the case of the pudding which arrives from nowhere depend on taking the mixing of ingredients as its cause merely in the way of a temporal predecessor? (Of course I mean here temporal predecessor of the pudding's appearance at this place and time: when the pudding begins to exist is precisely what is in question.)

But, first, it can still be the case in my example that a cause necessitates its effect. It might be that mixing ingredients in one's kitchen *necessitates* the arrival of the pudding from elsewhere in just the way I have supposed in the example. Second, I do not think Professor Anscombe is in a position to bring this objection. Although she thinks that Hume has not proved that there cannot be a logical connection between cause and effect, she believes that the claim that there is no logical connection between cause and effect 'seems to be almost universally true. I suspect indeed that it can be faulted only where it is tied up with his (Hume's) views on space and time' (p. 151). If this is how Miss Anscombe regards causation, there is all the less reason for her to object to my example.

---

[2] The quoted words are from a character in a dialogue: but the context makes clear that Miss Anscombe accepts this argument.

One might, though, reply to my example in this way. Even if we say we do not *see* the pudding being made of the ingredients in the sense that one can exclude, even as a bare possibility, the story of the arriving pudding, still we have a well-founded theory about how puddings are made. According to our ordinary picture of 'how things work', it is simply false that pudding-ingredients disappear when we are making a pudding and that puddings arrive mysteriously just subsequent to this disappearance. Besides, if someone doubts whether the ingredients in the completed pudding are the same as those in the pan, can't we devise tests that will answer the question?

But if we say this, why can't we say that, in the case of the uncaused effect, as far as any tests we have been able to devise show, no evidence exists to indicate that it arrived from elsewhere? That is to say, imagine that we have set up barriers, radiation-detectors, etc., aimed at showing whether an object is on the way and none arrives. If, under such circumstances, an object appeared to begin at a given place and time, why wouldn't we be justified in saying that we had reason to think it did begin there and then?

Whatever tests we devise will not be enough to eliminate any doubt. But to the extent that they are not, similar tests will leave room for doubt in cases where we claim to know the cause of something's coming into existence. If Professor Anscombe is right that we do not know whether an uncaused object has begun to exist at a particular time and place, we also do not know this about a caused object. Her argument proves too much.

In another sense, her argument does not prove enough. Suppose that we cannot know that an uncaused object that appears to begin to exist at a particular time and place really does so, since it might have come from elsewhere. This would in no way show that whenever it did begin to exist, even if it did not do so when it first appears to us, it had a cause. Miss Anscombe does not say that it does; but unless one supposes this, I do not see how her argument raises a difficulty for 'the serious supposition, that some object has come into existence without any cause' (p. 160). That we cannot tell whether something has come into existence without a cause is no reason against thinking that such a circumstance is possible.

6238 Orange Street,
Los Angeles, California 90048, U.S.A.

© DAVID GORDON 1984

THESEUS' CLOTHES-PIN

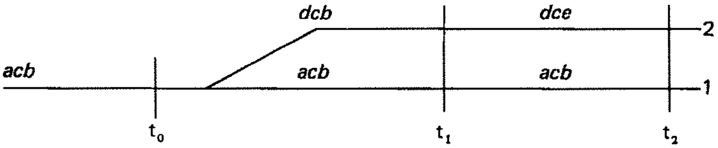
By HUGH S. CHANDLER

PLUTARCH reports:

> The ship wherein Theseus and the youth of Athens returned had thirty oars, and was preserved by the Athenians down even to the time of Demetrius Phalereus, for they took away the old planks as they decayed, putting in new and stronger timber in their place, in so much that this ship became a standing example among philosophers, for the logical question as to things that grow; one side holding that the ship remained the same, and the other contending that it was not the same. (Plutarch's *Lives*, 'Theseus')

This is the puzzle about Theseus' ship in its original (pre-Hobbesian) form. The same puzzle can arise in regard to a tree, a frog, or a person. It is a mystery to me why some philosophers say the puzzle only arises in regard to artefacts (see, for example, David Wiggins, *Sameness and Substance*, Blackwell 1980, pp. 86–99). Consider Theseus' tree. A golden-delicious apple tree suffers progressive rot. Bit by bit sections of its trunk, limbs, and even its roots, are removed and replaced by grafts from a Macintosh. Eventually there is so little golden-delicious left philosophers begin to argue about whether or not it's the same tree at all.

I'm going to concentrate on just one of the many issues Theseus' ship brings to light; but, for simplicity's sake, I turn the ship into a clothes-pin named 'Alice'. Alice is composed of two wooden legs, *a* and *b*, and a metal spring, *c*. In addition, there are two spare legs, *d* and *e* (the spring is not likely to wear out). I assume that our 'intuitions' about whether or not a clothes-pin can survive successive destruction and replacement of both legs (as opposed to the destruction and replacement of just one) are in conflict. Simplifying again, we have two possible sequences:



Along track 1 Alice suffers no loss of legs. But look at the sequence of events along track 2. First leg *a* is destroyed and Alice gets *d* as a replacement. Then leg *b* goes the way of leg *a* and is replaced by *e*. This means either that Alice no longer exists, or that Alice has two new legs.

The case helps us see problems about the *denotation* of the name 'Alice', and problems about its *intension*, or, if you prefer, the range of Alice's| 'counterparts'. I think Alice shows that ordinary proper names for things like ships, frogs and so forth, cannot be regarded as rigidly designating just one individual, or a cluster of individuals.

We are going to need two kind terms linked to the two obvious disambiguations of the *identity conditions* for clothes-pins. So let's say that a 'clothes-pin two' *can* survive successive destruction and replacement of both legs, while a 'clothes-pin one' *cannot* (that is to say, successive destruction of both legs *destroys* a clothes-pin one). And let's suppose that, in point of fact, Alice endures without any replacement of parts until after t_2. Then Alice is caught in a fire and reduced to ashes and a molten blob. Thus the world takes track 1.

The problem about the intension of the name 'Alice' is this: are both the clothes-pin one and the clothes-pin two on track 2 possible Alices (Alice counterparts), or is just one of these Alice? The problem about the denotation of the name is this: does 'Alice' (partially) denote *two* individuals, namely the clothes-pin one and the clothes-pin two on track 1, or does it denote just one individual there? (For an exposition of 'partial denotation' see Hartry Field, 'Quine and the Correspondence Theory', *The Philosophical Review*, April 1974, pp. 200-28.)

It would be absurd to hold that only the clothes-pin one or only the clothes-pin two on track 2 is Alice and *also* hold that 'Alice' partially denotes the clothes-pin one, and partially denotes the clothes-pin two, on track 1. Eliminating this absurdity, we have three candidate theories:

Theory A: Only the clothes-pin one or the clothes-pin two is the possible Alice on track 2; and 'Alice' denotes just the clothes-pin one, or the clothes-pin two, on track 1; but not both.

Theory B: Both the clothes-pin one and the clothes-pin two on track 2 are (partially) possible Alices; and 'Alice' denotes just one individual on track 1.

Theory C: Both the clothes-pin one and the clothes-pin two on track 2 are (partially) possible Alices; and 'Alice' (partially) denotes the clothes-pin one and (partially) denotes the clothes-pin two, on track 1.

It's fairly clear that theory A is unacceptable. What are we to suppose determines whether Alice is a clothes-pin one or a clothes-pin two? It cannot be actual future clothes-pin talk, for we may never bother to 'clarify' this matter (why on earth should we?). Are we to pretend that diligent research would — at some ideal limit — reveal that clothes-pins can, or cannot, survive replacement of both legs?

One might imagine that theory A is only implausible in regard to artefacts. But surely it fares no better as applied to trees or frogs? Future researchers may find it useful to stipulate that an individual frog can, or cannot, be said to exist after such and such replacement of parts. But then alternative stipulations would be equally 'correct'. Which of these possible future stipulations fixes the present denotation of a frog's proper name?

Consider theory B (I intend theory B to be one David Lewis could accept; see especially his Postscripts to 'Counterpart Theory and Quantified Modal Logic', *Philosophical Papers*, Vol. I, pp. 39–43). 'Alice' denotes just one individual with somewhat indeterminate potentialities. If 'Alice' denotes one individual, then, I take it, that individual must be *both* a clothes-pin one *and* a clothes-pin two. Suppose we name this clothes-pin one 'Alice one' and this clothes-pin two 'Alice two'. Presumably both of the new names, and 'Alice', all denote the same thing. Here we approach an inconsistent triad (see Saul Kripke, *Naming and Necessity*, e.g. p. 3; see also Keith Donnellan, 'Kripke and Putnam on Natural Kind Terms', *Knowledge and Mind*, Carl Ginet and Sydney Shoemaker editors).

(1) 'Alice one' and 'Alice two' rigidly designate their extensions.
(2) At t_2 on track 2 Alice one is not the same thing as Alice two (i.e. $\Diamond \sim$ Alice one $=$ Alice two).
(3) Alice one is (in fact, i.e. on track 1) the same thing as Alice two.

Claim (2) looks all right. Alice one no longer exists on track 2 at t_2. (Destruction of leg *d*, and, later, of leg *e*, destroy Alice one.) But Alice two still exists. Hence Alice one and Alice two cannot there and then be the same thing. Thus we must reject (1) or (3). Theory B, I take it, commits us to (3), and therefore to the rejection of (1).

The same point can be made in the language of counterpart theory. Theory B commits us to holding that 'Alice', 'Alice one', and 'Alice two', all denote the same thing. But Alice one and Alice two do not have the same counterparts on track 2. Hence theory B forces us to hold that one individual, *qua* Alice two, has a counterpart it lacks *qua* Alice one.

If 'Alice one', or 'Alice two', does not designate its extension rigidly, how can 'Alice' do so? Theory B yields that 'Alice' in fact denotes Alice one *and* Alice two. But it is not the case that wherever 'Alice' partially designates something it partially designates Alice one (remember track 2 at t_2). Hence, on this theory, 'Alice' does not rigidly designate its extension.

Let's shift to theory C. Here we suppose that 'Alice' partially denotes the clothes-pin one on track 1 and partially denotes the clothes-pin two there. The clothes-pin two survives (or has a counterpart) at t_2 on track 2; but the clothes-pin one does not. Proponents of this theory can hold that 'Alice one' and 'Alice two' are proper names rigidly designating their extensions.

What about 'Alice'? A name designates each of the things it partially designates. Thus 'Alice' designates both Alice one and Alice two on track 1. At t_0 and t_1 on track 2 'Alice' still designates Alice one and designates Alice two. What happens at t_2? 'Alice no longer exists' is there and then *true* in so far as 'Alice' is interpreted as 'Alice one', and *false* in so far as 'Alice' is interpreted as 'Alice two'. Hence (by the move Hartry Field recommends in such cases) 'Alice no longer exists' is neither true nor false there and then.

Does this theory allow 'Alice' to designate its extension rigidly? The theory says that 'Alice' does not designate just one individual, hence it certainly can't designate just one individual rigidly. But can the name rigidly designate *both* Alice one and Alice two? No; look again at t_2 on track 2. Thus 'Alice' does not rigidly designate each of the individuals it partially denotes.

In counterpart terms: Alice one and Alice two have different counterparts. Are we to think that Alice has a counterpart only where both Alice one and Alice two do? Clearly not, for, on that view, Alice does *not* have a counterpart at t_2 on track 2 (whereas, in fact, it is *indeterminate* whether or not Alice has a counterpart there then). Should we hold that Alice has a counterpart wherever either Alice one or Alice two has one? That won't do either since it would yield that Alice *could* survive successive destruction and replacement of both legs and this is indeterminate. The counterparts of Alice one and Alice two are Alice's *partial* counterparts.

In its generalized form, theory C will strike some readers as extravagant. C theorists presumably hold that an ordinary proper name partially denotes a distinct individual for each of its partial intensions. For many names, one guesses, the number of such intensions is non-denumerable.[1] The idea that ordinary names denote so many individuals may make theory C look less attractive than B. We naturally suppose that a frog, or a clothes-pin, is just one individual.

Summing up; theory A was declared unacceptable. Theory B may be all right; but it will not permit us to regard 'Alice' as rigidly designating its extension. Those who are fond of rigid designators must turn to C. 'Alice one' and 'Alice two' are rigid by hypothesis. But ordinary names like 'Alice' remain a disappointment.

University of Illinois at Urbana-Champaign, 810 South Wright Street,
Urbana, Illinois 61801, U.S.A.

© HUGH S. CHANDLER 1984

---

[1] List the infinity of non-actual possible worlds in which a candidate N is found, working from left to right. List vertically 'N''s partial intensions and assign a one or a zero in the column for each world indicating whether or not N is to be found in that world under that intension. The result is an infinite array of intensions; but an intension *not* in the array can be constructed by Cantor's diagonal procedure. Hence, the number of 'N''s partial intensions is non-denumerable.

PUTNAM'S BRAINS

By JANE MCINTYRE

IN *Reason, Truth and History* Putnam's announced intention is 'to break the strangle hold which a number of dichotomies appear to have on the thinking of both philosophers and laymen' (Hilary Putnam, *Reason, Truth and History*, Cambridge University Press 1981, p. ix; all references to Putnam are to this work). In wrestling with the old but vigorous dichotomy between objective and subjective theories of truth, Putnam appeals to his theory of reference to argue against a particular view of the relationship between mind and world – one in which the subjects of experience might be radically misled about the nature of the world and their relationship to it. Such a situation can be imagined by considering a brain in a vat, whose sensory experiences (if we choose to call them that), although qualitatively identical to our own, are produced and controlled by a machine. Putnam argues that the supposition that we might be brains in a vat is necessarily false. If successful, the argument would similarly establish that our relationship to the world cannot have certain features essential to that situation – that the world cannot be radically different from our best conception of it. Unfortunately, this interesting argument fails. In this paper I will examine two different versions of Putnam's argument. It is not my purpose here to criticize Putnam's theory of reference, but to analyse the arguments based on that theory.

Putnam rejects what he terms 'magical theories of reference' – theories that implicitly assume an intrinsic connection between words and the things they refer to. Reference, for Putnam, cannot be isolated from the abilities of language users. In particular, reference presupposes the ability to recognize and react to objects purportedly referred to when they are in the speaker's environment. Even a very sophisticated ability to construct and react appropriately to *sentences* would not, by itself, be sufficient to show that those sentences were *about* something, or had reference. As described by Putnam, a brain in a vat might have the latter sort of linguistic competence, but lacking interaction with (and therefore responsiveness to) ordinary objects, it could not refer to them. The terms in a language such brains might use ('vat-English') would not refer to the objects we refer to, even if their language duplicated ours in its entirety.

In the light of this view of reference Putnam argues that we could not be brains in a vat. The first version of this argument attempts to show that the supposition that we might be brains in a vat is self-refuting – that if we can consider whether it is true or false, it must be false. Putnam maintains that since brains in a vat cannot refer to (real) brains and (real) vats, they therefore cannot consider whether they are brains in vats. To consider whether I am a brain in a vat, I must be able to refer to brains and vats; but, on Putnam's account of reference, no brain in a vat can so refer. The heart of Putnam's argument, therefore, is that if I can consider whether I am a brain in a vat, then I am not one (ib. pp. 7–8).

The difficulty with this argument is that, given Putnam's account of reference, we have no reason to believe that we can actually consider whether we are brains in a vat. The mere use of the words 'I wonder whether we are brains in a vat' is not sufficient for considering whether we are. For, on Putnam's view, those words can be used (in vat-English) without referring to brains and vats; considering would require that you actually refer to those things. Putnam's first argument, by failing to show that we can consider whether we are brains in a vat, also fails to show that we are not brains in a vat.

Putnam presents a second argument in support of his conclusion. This version also rests on his theory of reference and the inability of brains in a vat to refer to actual brains and actual vats. On Putnam's view, a brain in a vat could only refer to *images* of brains and *images* of vats, even though it might use the same words with which we refer to brains and vats themselves.

In the light of this, Putnam argues:

> If we are brains in a vat, then the sentence 'we are brains in a vat' says something false... In short, if we are brains in a vat, then 'we are brains in a vat' is false. So it is (necessarily) false. (ib. p. 15)

This would be true, Putnam thinks, because the sentence 'we are brains in a vat' when uttered by a brain in a vat, would actually say that we are brains in a vat in the machine-created image, and this is false.

Putnam appears to take his argument to have the form 'p implies not-p, therefore not-p'. However, this is not actually the structure of the argument as stated. To arrive at an argument of this form we must assume the apparently innocent additional premise that if the sentence 'we are brains in a vat' is false, then we are *not* brains in a vat. Putnam's argument could then be stated:

(1) If we are brains in a vat, then the sentence 'we are brains in a vat' says something false (from p. 15, quoted above).
(2) If the sentence 'we are brains in a vat' says something false then we are not brains in a vat (implicit assumption).
(3) Therefore, if we are brains in a vat, then we are not brains in a vat (from (1) and (2)).
(4) Therefore, we are not brains in a vat (from (3)).

The additional assumption, captured above in (2), is not warranted, however, when we take into account that the false sentence is a sentence of vat-English. As I will show, Putnam's support for (1) requires that the false sentence be interpreted as a sentence of vat-English. But, once we do so, premise (2), although necessary to the argument, is no longer justified. Let us examine the reasoning behind premise (1).

According to Putnam, although a brain in a vat, speaking vat-English, can use the words 'I am a brain in a vat', what he refers to by using those words are sensory images of brains and vats. The property a brain in a vat attributes to itself when using vat-English is actually that of being an image of a brain in an image of a vat. The brain itself is not, of course, a sensory image, but according to Putnam, images are all it can talk about given its particular situation in the world. Since, by hypothesis, the brain is *not* an image, but the sentence 'I am a brain in a vat' would say, on Putnam's theory, that it is, the sentence is false. This is Putnam's reason for asserting premise (1).

But it is clear from this explanation that the falsity of 'I am a brain in a vat' derives from its status as a sentence of vat-English, and its consequence reference to brain images and vat images. The sentence is false because the brain is not an image. Therefore, the falsity of 'I am a brain in a vat' in *vat-English* does not entail that I am not a brain in a vat. The argument that supports premise (1) therefore defeats premise (2) and Putnam's argument is therefore unsound.

Cleveland State University,
Cleveland, Ohio 44115, U.S.A.

© JANE MCINTYRE 1984

AN ARGUMENT IN FAVOUR OF NON-CLASSICAL LOGIC FOR QUANTUM THEORY

By NICHOLAS HALLAM

IN his essay 'Against Quantum Logic' (ANALYSIS 43.2, March 1983) Jonathan Harrison criticises an argument, used in Professor Putnam's 'The Logic of Quantum Mechanics' (*Mathematics, Matter and Method*; Philosophical Papers: Volume 1), purporting to show the advisability of dropping the principle of distribution from classical logic.

Putnam's argument depends, according to Professor Harrison, upon the quantum mechanical principle that a particle cannot simultaneously have both a determinate position and a determinate velocity. Thus, if p_0 is a proposition ascribing to a particle a possible position and $v_0, v_1, \ldots, v_n, \ldots$ ascribe to it its possible velocities, then while the proposition

(1) $p_0 \cdot (v_0 \vee v_1 \vee \ldots \vee v_n \vee \ldots)$

may be true, and although the proposition

(2) $(p_0 \cdot v_0) \vee (p_0 \cdot v_1) \vee \ldots \vee (p_0 \cdot v_n) \vee \ldots$

is derivable from (1) using the classical distributive law of conjunction over disjunction, (2) must always be false. Putnam's conclusion is that the distributive law should be rejected.

Harrison believes this argument to be invalid. If quantum theory and the first conjunct of (1) are both true then the second conjunct

(3) $v_0 \vee v_1 \vee \ldots \vee v_n \vee \ldots$

is false. Consequently, in spite of Putnam's assertion to the contrary, (1) can only be true if quantum theory is incorrect, and so the argument for the adoption of a non-classical logic fails. Harrison states that he believes Putnam's error to arise from mistaking (3) for a logical truth; he, Harrison, agrees that if the particle is travelling with a determinate velocity then this velocity must belong to the set $\{v_0, v_1, \ldots, v_n, \ldots\}$, but he claims that there is another possibility which is not considered by Putnam, namely that, at the moment at which the particle may have ascribed to it a definite position, it has no determinate velocity. Indeed, he asserts, it is just this possibility that obtains when the first conjunct of (1) is true.

Harrison has presented what appears to be a plausible and persuasive analysis of Putnam's argument, but I believe that he has underestimated the case for quantum logic. The weakness of his account lies in its excessive appeal to the 'ordinary language' formulations of important features of quantum mechanics. The theory has a deep and highly sophisticated mathematical formalism; the fact that disputes about the physical interpretation of this formalism continue today, over fifty years after the development of quantum mechanics, bears witness to the difficulties involved in translating the theory into everyday language. If greater attention is paid to specific features of the mathematical formalism then the argument for the adoption of a non-classical logic may be seen in a more favourable light. By presenting an argument which is close, on my understanding, to Putnam's and which places greater emphasis on the formal mathematical aspects of quantum mechanics, more explicitly those of the Hilbert space formalism due to J. von Neumann, I hope to meet Harrison's objection.

In this formulation every physical system has associated with it a Hilbert space. A mapping is defined which takes physical propositions about the system to subspaces of this Hilbert space. For the sake of simplicity of exposition I shall assume, in what follows, that we are only interested in propositions about position and velocity, and that these magnitudes are nondegenerate.

Assume that $r_0, r_1, \ldots, r_n, \ldots$ are the position eigenvalues of a physical system S, and that $s_0, s_1, \ldots, s_n, \ldots$ are the velocity eigenvalues of the same system. The propositions p_i and v_j are introduced as abbreviations for 'the system S has position r_i' and 'the system S has velocity s_j', respectively. Then, for these basic physical propositions, the mapping S(P) is defined as follows:

$S(p_i) = V_{r_i}$ where V_{r_i} is the one-dimensional subspace of H(S) containing the eigenvector corresponding to the eigenvalue r_i,
$S(v_j) = V_{s_j}$ where V_{s_j} is the one-dimensional subspace of H(S) containing the eigenvector corresponding to the eigenvalue s_j.

This mapping is extended by recursion as follows:

$S(P \vee Q) =$ the span of the spaces S(P) and S(Q),
$S(P \cdot Q) =$ the intersection of the spaces S(P) and S(Q),
$S(\sim P) =$ the orthocomplement of the space S(P).

In a sense quantum mechanical events have been located in Hilbert space; on this basis truth conditions for the class of physical propositions about a system are prescribed as follows:

(4) P is true just in case the state vector of S lies in S(P).

The coordination of subspaces with propositions, and the truth conditions accompanying this mapping, accord well with our expectations where physical intuitions guided by popular expositions of quantum theory are strongest; e.g. for each i and j the proposition $p_i \cdot v_j$ is always false since the intersection of the two spaces V_{r_i} and V_{s_j} is empty – this is an empirical consequence of quantum mechanics. On the other hand these intuitions may mislead us. This applies to the proposition (3) which features so prominently in Harrison's account; this infinite disjunction will always be true since it is a consequence of an empirical assumption of quantum mechanics that the one-dimensional subspaces V_{s_1}, $V_{s_2}, \ldots, V_{s_n}, \ldots$ span the whole space.

We are now in a position to reconstruct the argument for the rejection of the principle of distribution.

Suppose that S is in the eigenstate of position corresponding to the eigenvalue r_0. Then proposition (1) is true: the state vector does lie in the intersection of V_{r_0} with the whole space. If we used the classical principle of distribution we could conclude that (2) is true. However, each proposition of the form $p_0 \cdot v_j$ corresponds to the empty subspace, and the intersection of the empty space with itself is empty, so we must regard (2) as false. Thus the contradiction which motivates the rejection of the distributive law has reemerged.

Of course Harrison is at liberty to reject the truth conditions (for physical propositions) captured in the formula (4), but it is unclear to me why he should feel this desirable. (4) is formulated explicitly in terms of the central concepts of the mathematical formalism of an experimentally successful theory. It provides a much more satisfactory criterion for the truthfulness of physical assertions than any criterion rooted in the vague notions of popular presentations of the theory.

There is, therefore, a stronger case for the adoption of a non-classical logic for quantum theory than Harrison apparently accepts, but there are two objections which may be raised. The first is that any change in logic is a philosophically uninteresting matter amounting to a mere relabelling of concepts. This may be true of some changes (e.g. that effected by interchanging 'and' and 'or') but clearly does not apply here; no amount of relabelling can avoid the contradiction that results from the use of the distributive law with the non-classical truth conditions. The second objection is that the rationality of any enquiry depends upon classical logic and the accompanying classical concept of truth. Without attempting to deal with this fully I shall suggest two ways in which one might try to reply to this objection. One approach may be to take the notion of rigid designation, originated by Saul Kripke and used by Putnam in his papers on language, and extend it in some fashion to the logical constants. It may then be possible to argue that there is no meaning change involved in the transition to the non-classical logic, but that the transition is part of a process of approaching to the, as yet not fully comprehended, meanings of the logical constants. Another way of replying to the last objection may be to argue that it is fruitful to consider extending the holistic conception of science to embrace the laws of logic. If this is so then the adherence to a single concept of truth and a single logic may appear inimical to successful enquiry rather than as a *sine qua non* for it.

Perhaps there are good reasons for holding classical logic unrevisable in principle. But if not, then empirical considerations may make the adoption of a non-classical logic for quantum theory advisable.

University of Sussex, Falmer,
Brighton, Sussex BN1 9QX

© NICHOLAS HALLAM 1984

WHY THE DISTRIBUTIVE LAW IS SOMETIMES FALSE

By P. F. GIBBINS

THE distributive law is sometimes false in quantum logic because of the *superposition principle* (and the projection postulate) in quantum mechanics. When a quantum proposition (in van Fraassen's 'elementary language of quantum mechanics') is true of a quantum system in each of two states, it is true when that system is in any superposition of those states. If 'p' is true of a quantum system when it is in state 1 and 'q' true of it when it is state 2 then '$p \vee q$' is true of it when it is any superposition of states 1 and 2, even though neither 'p' nor 'q' may be true of it when it is in that superposition, a fact which is essentially equivalent to the failure of distributivity (see for example C. Piron 'Foundations of Quantum Physics', Benjamin, 1976, p. 21). Clearly, quantum logical 'or' behaves non-classically.

This superposing of states has no physically significant counterpart in classical mechanics, which is why the logic of classical mechanics is classical and why the logic of quantum mechanics is quantal. Consider the superposition of two orthogonal quantum states of a system, each (for example) corresponding to the system's having its position in a certain range which is disjoint from that of the other. A measurement of position will result in the system being found in one of the two disjoint ranges (an outcome which follows from the projection postulate). In classical mechanics two position-vectors which sum to the position-vector of a classical particle can be 'superposed' (mathematically) but a position measurement will not show that the particle's position is necessarily given by one of the two vectors. One can add states in classical mechanics but this adding has no physical significance. In quantum mechanics the *projection postulate* gives the principle of superposition its physical significance.

Try to force the classical connective 'or' on the elementary language of quantum mechanics and you must deny that the truth of a pair of quantum propositions is preserved under superpositions. This is why we must adopt quantum logic for that language. To be faithful to the superposition principle we must say that *the classical disjunction of two quantum propositions is in general not a quantum proposition at all.* The failure of distributivity in quantum logic – which simply is the logic of the elementary language of quantum mechanics – comes first. The uses to which philosophers (like Putnam) have put that failure come second. This much should be uncontroversial. What is controversial is how quantum logic helps (if it does help) to resolve the paradoxes of quantum mechanics.

Now contrast the account given by Professor Harrison (ANALYSIS 43.2, March 1983, pp. 83–5) who voices the plain philosopher's discontent with the failure of the distributive law in quantum logic as it is presented by that historical figure, the Hilary Putnam of 'The Logic of Quantum Mechanics' (LQM) (*Mathematics, Matter and Method* Vol. 1, Cambridge, 1979, pp. 174–97). To Harrison, reviewing the philosophical literature, it must seem that quantum logic is a philosopher's aberration. With a slight modification, the argument (found in LQM) which offends Harrison is as follows.

Imagine a quantum system which has its spatial location restricted to a finite volume of physical space. The conjunction of

(i) the proposition p_0 asserting that the quantum system is actually in this volume, with

(ii) the exhaustive, infinite disjunction $[\vee v_i]$ of all the propositions which assert that its momentum lies in some finite volume of momentum space

(1) $p_0 \wedge [\vee v_i]$

is true. But the distributive expansion

(2) $\vee [p_0 \wedge v_i]$

of (1) is false (indeed a quantum logical contradiction) since each of the disjuncts in (2) is a quantum logical contradiction. (Even in quantum logic a disjunction of contradictions is a contradiction.) Hence the classical distributive law is false.

Harrison complains that he had always supposed that the original conjunction (1) is false because the second conjunct $[\vee v_i]$ is false. ('Given that a particle's position is determinate, it is false that it has any of the velocities specified in the second conjunct of (1)'.) But Harrison's supposition is just what Putnam wanted to reject. The Putnam of LQM was motivated by quantum mechanical realism, by the collapse of wave-particle duality in favour of the particle, by the idea that a quantum system takes simultaneous exact values for all its dynamical variables. He *used* (rather than argued for) the well-known failure of the distributive law in quantum logic to *defend realism against the paradoxes* (a task in which, incidentally, it fails). So Harrison's supposition can cut no ice with (the historical) Putnam.

I should agree that Harrison's diagnosis of what he takes to be Putnam's error – that Putnam thinks that $[\vee v_i]$ is a logical truth (which it is) because a quantum system must have some definite momentum (which it need not) and because the disjunction exhausts all the momenta there are – may explain Putnam's having the untenable idea of quantum logical realism. But the failure of the distributive law in quantum logic has a more substantial source, namely quantum theory itself. Ironically for Putnam's former quantum logical realism, one simple and elegant way of expressing the non-classical nature of quantum logic is to say (as we did say) that a quantum logical disjunction may be true even though neither of its disjuncts is true. This is bad news for quantum logical realism which has it that necessarily some one of the v_i's in $[\vee v_i]$ is true. All this leaves untouched the really interesting question. What role should quantum logic play in our understanding of the quantum world? The merit of Putnam's LQM is that it raised in a new way the question of the freedom the theoretical physicist and the metaphysician have (here the distinction is blurred) to adopt a non-classical logic.

One final point. For reasons technical, though also relevant to the question of quantum mechanical realism, it would be better if philosophers dropped Putnam's example of the move from (1) to (2) above with each v_i asserting an exact momentum. Professor Harrison is not to be blamed for adopting it; almost everyone does. But notice that the disjunction $[\bigvee v_i]$ is non-denumerable. And as a matter of quantum fact, one cannot shrink the volume elements in the v_i to zero. Strictly speaking, a quantum system cannot have exact values for *either* its position *or* its momentum (pace Dirac and the delta-function).

Programming Research Group,
Computing Laboratory,
8-11 Keble Road, Oxford

© P. F. GIBBINS 1984

A PUZZLE ABOUT LOGICAL EQUIVALENCE

By STOJAN T. DOBROSAVLJEVIĆ

QUITE often, politicians make statements like the following:

(i) Those who are not with us are against us.
(ii) Those who are not against us are with us.[1]

Statement (i) would usually convey that the speaker does not want to compromise with those who do not belong to his party: these persons should be treated as enemies, and one should not tolerate them. On the other hand, the tone of (ii) is conciliatory: all those who are not overtly fighting the party of the speaker should be treated as friends, and one should tolerate them. So, statements (i) and (ii) usually convey two very different messages, and they cannot be interchanged.

It is then rather surprising to find that statements (i) and (ii) are logically equivalent. Let A be the set of those who are with us, and B the set of those who are against us. Furthermore, let $\bar{X}$ be the complement of X, and let $\subseteq$, $\cap$, $\cup$ and ϕ be defined as usual. Then statements (i) and (ii) can be represented as follows:

---

[1] Both of these statements can be derived from the Gospel: (i) should be compared with 'He that is not with me is against me...' (Matthew 12:30 and Luke 11:23), and (ii) should be compared with 'For he that is not against us is on our part' (Mark 9:40) and '... for he that is not against us is for us' (Luke 9:50).

(i) $\bar{A} \subseteq B$
(ii) $\bar{B} \subseteq A$

and it is not difficult to show that in classical logic these two statements are logically equivalent.

How is it possible that (i) and (ii), in spite of being logically equivalent, convey two messages which are so different?

This puzzle can be solved as follows. A quite natural assumption about A and B is that

(*) There are individuals which are neither in A nor in B (i.e. $\overline{A \cup B} \neq \phi$).

The solution of our puzzle hinges on this assumption. Both (i) and (ii) contradict assumption (*), but not in the same way. Let C be the set of those individuals who are neither in A nor in B. An ordinary person, assuming (*), will assume that C is nonempty. But now, (i) suggests that members of C will henceforth be counted as members of B, which presumably means that they will be treated as enemies. On the other hand, (ii) suggests that members of C will henceforth be counted as members of A, which presumably means that they will be treated as friends. This explains the intolerant tone of (i) and the conciliatory tone of (ii).

To see that this is indeed the right solution of our puzzle consider cases where it is quite clear that A and B exhaust the whole domain of individuals, i.e. $\overline{A \cup B} = \phi$. For example, A can be the set of all men, and B the set of all women. Then $\bar{A} \subseteq B$ and $\bar{B} \subseteq A$ sound tautologous, and they both have the same neutral tone. It is exactly where (*) is assumed that $\bar{A} \subseteq B$ and $\bar{B} \subseteq A$ cease to sound tautologous and succeed in conveying something.

Note that it is irrelevant for this solution whether A and B are disjoint or not. This is because (i) and (ii) are not referring to those who might be both A and B in the way they are referring to those who are neither. There is nothing in (i) and (ii) to contradict the natural assumption that A and B are disjoint.

Let us now set aside a problem which might arise from the fact that in proving the logical equivalence of (i) and (ii) we use *classical* logic. In intuitionistic logic the law of contra position

$$(\neg p \rightarrow q) \rightarrow (\neg q \rightarrow p)$$

is not generally valid, and since the classical equivalence of (i) and (ii) is presumably based on this law, it might happen that we are unable to show that (i) and (ii) are logically equivalent in intuitionistic logic. There seems to be no good reason to assume that in contexts where (i) and (ii) are used we should rely on intuitionistic rather than classical logic. But apart from that, it is possible to construct puzzles analogous to the puzzle with (i) and (ii) where logical equivalence satisfies intuitionistic standards. Consider the following statements:

(iii) Good men are not our enemies.
(iv) Our enemies are not good men.

If A is the set of good men, and B the set of our enemies, then these statements are represented by:

(iii) $A \subseteq \bar{B}$
(iv) $B \subseteq \bar{A}$

and it is clear that (iii) and (iv) are logically equivalent even according to intuitionistic lights. The tone of (iii) may well be taken as conciliatory, and the tone of (iv) as intolerant. This is because we may assume that

(**) A is not disjoint from B (i.e. $A \cap B \neq \phi$).

For let D be the set of those who are both good men and our enemies. According to (**) this set is nonempty. But now, (iii) suggests that members of D will henceforth not be counted as enemies, whereas (iv) suggests that members of D will henceforth not be counted as good men.

It should be clear by now that on the pattern of these puzzles we can construct many analogous puzzles based on assumptions like (*) and (**). These assumptions will state the nonemptiness of sets like $\overline{A \cup B}$ and $A \cap B$, or possibly something else, and logically equivalent statements will run counter to these assumptions in different ways.

Cara Lazara 11,
11000 Belgrade, Yugoslavia

© STOJAN T. DOBROSAVLJEVIĆ 1984

WRIGHT'S REPLY TO BENACERRAF

By MARIE MCGINN

PAUL BENACERRAF has objected to a platonistic conception of numbers on the ground of an essential indeterminacy of reference of mathematical terms ('What numbers could not be', *Philosophical Review* 74, 1965, pp. 47–73). He argues that since no single reduction of number-theory to set-theory can be established as uniquely correct, there is no answer to the question which sets the numbers really are. And he goes on:

> If numbers are sets, then they must be *particular* sets, for each set is some particular set. But if the number 3 is really one set rather than another, it must be possible to give some cogent reason for thinking so; for the position that this is an unknowable truth is hardly tenable. But there seems to be little to choose among the accounts. Relative to our purposes in giving an account of these matters, one will do as well as another, stylistic preferences aside. There is no way connected with the reference of number words that will allow us to choose among them, *for the accounts differ at places where there is no connection whatever between features of the accounts and our uses of the words in question.* If all the above is cogent, then there is little to conclude except that any feature of an account that identifies 3 with a set is a superfluous one – and that therefore 3, and its fellow numbers, could not be sets at all. (op. cit. p. 62)

Crispin Wright has attempted to remove the antiplatonist sting from Benacerraf's argument, by insisting that it represents nothing more than a particular case of Quinean indeterminacy (*Frege's Conception of Numbers as Objects*, Scots Philosophical Monographs no. 2, 1983). And Quinean arguments for indeterminacy in the case of ordinary referential expressions have not in general led us, Wright argues, to conclude that none of our ordinary singular terms is genuinely referential. Why, therefore, should it undermine our confidence in the notion of reference in the mathematical case? Wright sums up his argument as follows:

> ... Benacerraf has not succeeded in developing a specific argument against number-theoretic platonism. He has a good argument for the inscrutability of reference of numerical singular terms with respect to the domain of classes; but he fails to honour the fact that it sits in the company of powerful arguments for the inscrutability of reference generally. There is nothing especially difficult about numerical terms in this respect. In a nutshell: we have a powerful case for the indeterminacy of numerical language into the language of classes. But we ought to draw Benacerraf's antiplatonist conclusion only if, confronting similar indeterminacies with respect to the language of classes itself, or our talk of rabbits, we are prepared to infer similarly that singular reference to classes, or to rabbits, is illusory. (ibid. p. 127)

I think the attempt to assimilate Benacerraf's argument to Quine's, and hence to reduce the sense of there being a *special* difficulty about the reference of number terms, is mistaken. There is, for one thing, the following striking difference between the two arguments.

Quine's argument for the indeterminacy of translation of the native word 'gavagai' into English turns upon the existence of an exactly parallel indeterminacy of translation of the native 'auxiliaries of objective reference', i.e. of the articles, the pronouns, the singular and plural, and, in particular, the identity predicate (v. *Word and Object*, MIT 1973, chapter 2 *passim*, but e.g. pp. 53–4). Thus, the possibility of translating 'gavagai' by any one of the English terms 'rabbit', 'rabbit stage', 'undetached rabbit part', etc., depends upon the possibility of our compensating in our translation of the apparatus of individuation so as to produce translations of whole sentences that capture the natives' patterns of assent and dissent. For example, the possible translations of 'gavagai' mentioned above require us to juggle the translation of the native identity predicate in such a way that the following translations of whole statements of identity are, respectively, the result: 'a is the same rabbit as b', 'a is a stage of the same rabbit as b', and 'a is an undetached part of the same rabbit as b'.

By contrast, adoption of either the Zermelo or the von Neumann definitions of number in terms of sets requires no reinterpretation of the numerical sign of identity. Thus, all that varies in the following statements of identity:

(i) $3 = [\emptyset, [\emptyset], [\emptyset, [\emptyset]]]$
(ii) $3 = [[[\emptyset]]]$

is the term that occurs on the right hand side; there is no need for compensatory reinterpretation of '='. Thus, in the mathematical case, the indeterminacy appears to touch not, as it were, the language of mathematics as a whole, but only a specific part of it, namely the singular terms. We can hold the apparatus of individuation constant while altering the objects assigned as reference to the individual terms. And this is, clearly, a much more serious indictment of the idea that number terms are genuinely referential (or, at any rate, of the idea that they refer to *the numbers*) than an entirely general argument for Quinean indeterminacy: there *is* something special about the mathematical case.

The difference between the two sorts of case is again brought out by considering what sort of response we might make to each of the two indeterminacy arguments. Suppose for the sake of argument that we accept Gareth Evans's claim ('Identity and Predication', *Journal of Philosophy* 72, 1975, pp. 343-64) that, by adding to Quinean constraints on translation constraints that are motivated by the need to provide a plausible semantics for the native language, we remove all indeterminacy surrounding the interpretation of 'gavagai'. It is clear, I think, that no analogous suggestion will help at all in the mathematical case: Evans's manoeuvres are not even relevant to the question of how we might eliminate all but one of the set-theoretic treatments of number-theory.

How might we, then, respond to Benacerraf's argument? Plainly, there can, in this case, be no thought of *removing* the indeterminacy at all: we must somehow learn to live with it. The question, therefore, must be: what does the sort of indeterminacy that Benacerraf draws attention to show? I think that one thing it shows is that the notion of *the numbers* has no real role to play in arithmetic. It is not, in other words, that Benacerraf establishes that numbers are not sets, but that in doing number-theory we need not commit ourselves to the existence of a special sequence of abstract objects, *the numbers*, but only to some one of a large number of sequences of *sets*. Benacerraf has not, therefore, shown us how to escape commitment to abstract objects altogether; he has only made it clear that the abstract objects we must commit ourselves to are sets, rather than numbers, and that we have a choice about just which sets we say there are (this, of course, is essentially Quine's view). And there is plainly little cause for comfort to the anti-platonist in that.

University of York,
Heslington, York YO1 5DD

© Marie McGinn 1984

WHY PHYSICS CAN'T BE NOMINALIZED

By Alan McMichael

Assuming the adequacy of Quine's criterion of ontological commitment, nominalism is a tenable doctrine only if we can formulate our scientific theories, and in particular our physics, without using quantifiers that range over abstract entities. One problem for nominalism is that classical mathematical theory resists such formulation. Nominalists have tried to skirt this difficulty, arguing that not all of classical mathematics need be given a nominalistic paraphrase, only those portions that are necessary for physical science. Even those portions are quite formidable, however, so formidable that Hilary Putnam has declared the nominalists' situation hopeless:

> If the numericalization of physical magnitudes is to make sense, we must accept such notions as function and real number; and these are just the notions the nominalist rejects. (*The Philosophy of Logic*, Harper & Row 1971, p. 43)

Actually, things are not quite so bad as Putnam suggests. Hartry Field's recent book, *Science Without Numbers* (Princeton 1980), contains some ingenious ideas for solving the mathematical difficulties. Field argues, successfully I believe, that physical science can jettison the specially mathematical apparatus of real numbers and associated functions. This can be done because the correlation of physical quantities with real numbers is simply a roundabout means for describing a structure that is actually *intrinsic* to those quantities. Of course, since physical quantities can be quite conveniently thought of as properties and relations, such as *weighing ten pounds* or *being one metre apart*, this result does not on its own support the nominalists' case. But Field goes further, arguing that relationships between physical quantities reduce to relationships between concrete things. (For example, the statement, 'The distance of X from Y is twice the distance of U from V', can be paraphrased by, 'The line segment XY consists of disjoint segments S and T, each of which is congruent to the segment UV'.) This argument is less safe than the first, since it involves some controversial assumptions. For example, Field assumes the existence of an actual continuous manifold of space-time points.

I shall not be concerned, however, with the details of Field's proposal. Instead, I shall argue that in attempting to solve the mathematical difficulties, Field and other nominalists have ignored more elementary objections. They have tried to show that mathematically complex relationships between physical quantities can be expressed without quantification over abstract objects, and in this endeavour, I agree, they have met with some success. But they have overlooked the fact that physical theories not only tell us how given physical quantities are related, but also *what physical quantities there are*. Since physical quantities are properties and relations, this means that our scientific theories tell us that *there exist certain properties and relations*. Unlike those statements of relationships between physical quantities which can somewhat adequately be reformulated in terms of relationships among individuals, these simple existence assertions do not admit of nominalistic paraphrase. (It is a fallacy of division to suppose that if a nominalist can succeed in paraphrasing statements expressing relationships between universals, then he can also successfully paraphrase statements expressing what, in the view of the realist, those relationships *entail*, namely the existence of the universals so related. Consider that '*Longer than* is a transitive relation' has the nominalistically acceptable paraphrase 'For any x, y and z, if x is longer than y and y is longer than z, then x is longer than z', but that it is not at all clear from this how one is to paraphrase the weaker assertion 'There exists a transitive relation'.)

As an illustration, consider Field's own formulation of the Newtonian theory of space and time. There are three basic predicates: a betweenness predicate, a 2-place simultaneity predicate, and a 4-place spatial congruence predicate. What I claim is that Field has *not* succeeded in rewriting the Newtonian theory of space and time, *except* under the assumption that his formulation implicitly entails the existence of relations corresponding to those predicates, or at least the existence of a set of relations corresponding to some adequate set of predicates definable in terms of Field's basic predicates. If, however, his formulation is taken to have this implicit entailment, it is not a truly nominalistic formulation.

Why must we suppose that the Newtonian theory of space and time entails the existence of certain relations? I take it that this is one of the features which distinguishes the Newtonian theory from others. For example, one respect in which it differs from Einstein's theory is the postulation of a 2-place relation of temporal simultaneity. Einstein's theory does not postulate such a relation; in fact, it denies the existence of absolute simultaneity. The closest analogue in Einstein's theory is a 3-place relation of simultaneity-relative-to-an-observer. To all appearances, then, these two theories disagree on the question of whether a certain sort of relation exists. My challenge to nominalists is to ask them to produce some other construal of the disagreement. In what other terms can the disagreement be framed?

There are other differences between the two theories, of course. According to Einstein's theory, spacetime may possess an *intrinsic curvature*, and it may even possess a *different topological structure* from that of Newtonian spacetime (Lawrence Sklar's *Space, Time and Spacetime*, University of California Press 1974, gives a good informal description of these differences). In order, therefore, to isolate the difference regarding simultaneity, let us concentrate on what the two theories say about those *regions* of spacetime in which there is no pronounced curvature and which are topologically normal. This concentration on what we might call *normal local spacetime* is not entirely *ad hoc*. It is a matter of historical fact that the portion of Einstein's theory that deals with normal local spacetime, Einstein's *special* theory of relativity, was first to receive experimental support.

Suppose we try to cast the Newtonian theory and Einstein's theory into first-order form, in the manner suggested by Field. The two theories will yield identical consequences concerning the topology of normal local spacetime, but will differ in their metrical pronouncements. The Newtonian theory will assert that events are governed by ordinary metrical laws for distance in space and duration in time. The Einsteinian theory will assert only that events are ordered by a more general set of metrical laws, the laws of the Minkowskian 'pseudo-metric'. Indeed, regarding the geometry of normal local spacetime, and ignoring 'bridge laws' specifying how particular geometrical quantities are to be measured,[1] we will find that this is the *only* difference between the two theories: the Newtonian theory postulates additional metrical relationships, such as those involving absolute temporal quantities. But since the Einsteinian theory is only first-order, it can contain no *denial* of the existence of the additional metrical quantities nor, therefore, any denial of the Newtonian laws in which they appear. Consequently, the two first-order theories yield *logically compatible* views of the *structure* of normal local spacetime.

In connection with bridge laws, however, the compatibility disappears. The Newtonian theory plus its bridge laws entails false consequences for observation – consequences, for example, asserting the consistency of measurements of simultaneity for different observers. But if this is the *entire* extent of the disagreement of the two theories concerning the nature of normal local spacetime, then we are left with a puzzling question: Given the possibility of modifying the Newtonian bridge laws, so that these false consequences for observation do not arise, why should anyone assert, as physicists do, that the Newtonian theory gives a false view of normal local spacetime and opt for Einstein's theory instead?

---

[1] I do not claim that there is an interesting *general* distinction to be made between the 'bridge laws' of theories and their 'structural laws'. What I am alluding to in the present case is a purely conventional distinction, one to which I attach no great philosophical significance.

The problem with the usual Newtonian bridge laws is that they assume that absolute simultaneity can be measured, by normal methods, in any reference frame. To avoid inconsistency, it is necessary to suppose that absolute simultaneity can be correctly measured, in the normal way, only in a single 'principal' reference frame and that, in other frames, a correct judgement of simultaneity can be deduced only by assuming that measuring rods are shortened and clocks run slow. It is possible to write this new stipulation into the bridge laws. In fact, Lorentz and other scientists were engaged in just such a project at the turn of the century (not for the purpose of saving the notion of absolute simultaneity, but for the related purpose of saving the notion of absolute rest in the luminiferous ether!). However, in the total theory that results from this enterprise, there exists no good means for determining what the 'principal' reference frame is. Obviously it won't do simply to *stipulate* the principal reference frame. That would amount to an admission that there is no *objectively correct* principal reference frame and, consequently, that the notion of simultaneity is not absolute. The alternative, however, is to suppose that there is an objectively correct reference frame, but that there exists no means for determining what it is. In this scheme, the simultaneity predicate is completely idle.

This sort of theory can be rejected, as Lorentz's whole approach *was* rejected,[2] on grounds of economy. But I think that the economy that is desired is not merely an economy of symbols. Although the first-order formulation of the Newtonian theory contains no explicit assertion of the existence of the simultaneity relation, a full second-order formulation would. The economy we seek concerns that second-order fomulation: We do not think that we should postulate a relation of absolute simultaneity between events since, not only do we lack evidence that events stand in that relation, but, more importantly, we lack any criterion for determining when they do. In the absence of such a criterion, we lack any grounds for supposing that such a relation exists.

[2] See Kenneth Schaffner's article 'Space and Time in Lorentz, Poincaré and Einstein', in *Motion and Time and Space and Matter*, ed. by Peter Machamer and Robert Turnbull (Ohio State University Press, 1976). Einstein criticized Lorentz's views as *ad hoc* and artificial. Poincaré thought that Lorentz required 'too many hypotheses'.

Of course, nominalists are in a position to maintain that the first-order Newtonian theory ought to be rejected on pragmatic grounds, due to its overly complex set of basic predicates. The problem is that they seem to have no basis on which to maintain that the Newtonian theory is *false*. This problem can be evaded if nominalism is rejected, for we may then say that the full formulation of the Newtonian theory is a second-order formulation, a formulation that contains a false assertion of the existence of absolute simultaneity.

Assertions of the existence of basic relations are not the only scientific claims that are troublesome for nominalists. Some theories assert the 'completeness', in some sense, of their primitive notions, and these completeness assertions are universal generalizations about properties and relations. Descartes made such a claim when he stated that 'the entire nature of body consists in extension alone', by which he presumably meant that all the intrinsic properties of bodies and all the relations which hold solely among them are *geometrical* properties and relations. A modern scientist makes such a claim if he says that such-and-such a list is a complete list of all the fundamental kinds of particles. Such claims are notoriously apt to be false, but that is no comfort to the nominalist, for their denials, which are in fact existence assertions, are then true, and they are just as resistant to nominalistic paraphrase.

Faced with this problem, a nominalist might choose to concede that physical theories cannot be satisfactorily formulated without quantification over relations but then argue that this entails no extension of his basic ontology, on the ground that relations are not *substances*, and they do not exist *independently* of individuals. Such a response, however, would simply be a rejection of the conditions of the problem, for I took it for granted at the outset that a nominalist paraphrase of a physical theory should be judged by *Quine's* standard of ontological commitment. It is precisely because I need only show the necessity for *quantification* over relations that I am confident that *this* nominalist project cannot succeed. Whether a nominalistic paraphrase is possible when judged by some other standard is not here at issue.

Having stated the difficulty for my nominalist opponents, let me now address a question that will occur to my realist friends. I have said that the theories of Newton and Einstein differ concerning the existence of a certain relation. Since I accept the conventional view that each of these theories is logically possible, I must also maintain that relations can have contingent existence. My view is therefore like the *empirical realism* of D. M. Armstrong,[3] and unlike platonic realism. According to platonic realism it is a necessary truth that if a relation exists, its existence is logically necessary.

---

[3] *Universals and Scientific Realism* (Cambridge University Press, 1978). I have not assumed, as some realists do, that a universal must have instances in order to exist. However, I suspect that these realists would agree with much that I have said.

The departure from platonic realism should not be overstated. On my theory, the existence of relations is, in a certain sense, necessary existence, although not logically necessary existence. Relations exist eternally and their existence is a matter of natural law. Moreover, it is compatible with my view that the existence of some relations, such as identity and the relations of pure mathematics, is logically necessary.

But I cannot bring my views into any closer agreement with the platonic position, because I know of no compelling argument for the supposition that all relations exist of logical necessity. Indeed, the supposition strikes me as too fantastic. For example, I have heard that a certain author who writes about religious experience claims that there are actually four planes of reality, only one of which we inhabit in our ordinary waking lives (Huston Smith, *Forgotten Truths*, Harper & Row 1977). His claim is logically possible. I suppose, therefore, that in addition to the relations of ordinary space and time, the relations that Newton and Einstein talked about, the platonic realists must admit that there exist other basic relations, the ones that this fellow supposes to exist among his four planes of reality. Maybe they judge this admission harmless, since they can claim that although these relations exist, they are of no great consequence, for there is only one plane of reality to bear the relations. But even if the admission is harmless, in the sense of being logically coherent, it does offend our sense of economy and befog our sense of the distinction between fiction and reality. Some good reason should be given before we accept it.

Perhaps the reason many philosophers believe that the existence of properties and relations is logically necessary is that these entities seem to provide the very framework within which judgements of logical necessity are made. It is hard to see, however, how this impression can give rise to an argument. It is a necessary truth, for example, that nothing can be both red and green in the same place at the same time, but this fact can be easily squared with the view that the existence of the properties *red* and *green* is logically contingent: There are possible worlds in which these properties exist, and in those worlds it is a matter of natural law that these properties exclude one another; in possible worlds in which these properties do not exist, nothing, *ipso facto*, can possess both properties.

We have adequate means for the description of logical possibilities even if we do not assume that all logically possible relations and properties actually exist. A logical possibility can be described in terms of the notions of pure mathematics and in terms of such empirical notions as do exist, without assuming the existence of *all* the empirical properties and relations which would be necessary to actualize that logical possibility, just as Pegasus can be described in terms of empirical properties and relations and without assuming his existence. It is in this way that we can, for example, describe the possible existence of the relation of absolute simultaneity. Absolute simultaneity is a relation such that, if it were to exist, it would be an equivalence relation among events, and it could be measured in the way that physicists once supposed. It is because no relation meets this description that we can conclude that no such relation in fact exists.

To platonic realists who are not convinced by these considerations, I may offer the same challenge I posed to nominalists: Certain disputes in science seem to be disputes concerning the existence of certain relations. How else are these disputes to be understood? If platonic realists are right, relations exist of logical necessity, and it is difficult to see how their existence can be a subject of empirical investigation.

It is possible that platonic realists, who have more devices at their disposal than nominalists, will be able to meet this challenge. I believe that any interpretation they may offer, however, will be less straightforward, and less plausible, than the one I have advocated.[4]

Virginia Polytechnic Institute
and State University, Blacksburg,
Virginia 24061, U.S.A.

© ALAN MCMICHAEL 1984

---

[4] I am grateful to the referee for his helpful suggestions.

McGINN ON EXPERIENCE OF PRIMARY AND SECONDARY QUALITIES

By DAVID MCNAUGHTON

COLIN McGINN has recently argued for what he calls an inseparability thesis (IT) about perception: 'for any actual and possible sense the content of experiences delivered by that sense must be both as of secondary qualities and as of primary qualities' (*The Subjective View*, Oxford 1983, p. 81). I shall contend that the inseparability thesis does not obtain universally; there can be sense experience which is as of secondary qualities alone and sense experience which is as of primary qualities alone.

Of the five traditional sense modalities IT seems to be clearly true only of sight; it is, as McGinn insists, impossible to imagine what it would be like to see a world of objects possessing primary qualities but yet completely colourless. (I am here following McGinn in supposing that it is an *a priori* truth that colours are secondary qualities. If it were an empirical discovery that colours are secondary qualities then IT might have been false even of sight.) It is, perhaps, not quite so obvious that there could not be an experience of colour which was not as of any primary quality (McGinn does not discuss this point). Might there not be a creature whose visual experience was solely of expanses of colour with no boundaries and hence, it may be suggested, with no size or shape? However, if there could be such a perceiver he would have experience as of colour patches possessing the primary quality of extension, even if they were not conceived of as having *determinate* size or shape.

Are there sense modalities in which it is possible to have experiences as of secondary qualities without having experiences of primary qualities? The clearest example is the sense of smell. Indeed, far from its being the case that one cannot conceive of an olfactory experience which is not also of primary qualities I find it hard to imagine what primary quality could be detected by that sense alone. I suggest that the same is true of hearing and, perhaps more controversially, of taste.

Can there be experiences which is as of primary but not as of secondary qualities? It is certainly the case that tactile experience in humans is as of both sorts of quality. However, IT is a claim about possible as well as actual experience. I suggest that we can conceive of a form of tactile experience in which we are aware only of the primary qualities of objects. Imagine a man who has no sensory receptors at the surface of his body so that he has no sensations of felt temperature, texture, etc. (compare Hume's man with a palsied hand, *A Treatise of Human Nature*, ed. P. Nidditch, Oxford 1978, p. 230). There seems no reason to believe that this man could not, like us, be aware of the position of his limbs without the use of sight. He would then be able to detect the size, shape and solidity of, say, a wooden cube, simply by exploring its surface with his hands. In so doing he would not be aware of any of the secondary qualities, such as heat or cold, which would normally be involved in the tactile experience.

McGinn suggests (p. 85, footnote 21) that tactile experience always involves felt qualities relating to hardness, softness, pressure, etc. But we need not suppose that our observer is aware in detecting the difference between, say, a tennis ball and a squash ball, of anything other than the degree of effort required to produce compression of the ball. (It may help here to think of what such an experience would be like for a man who has an articulated artificial hand.)

It may be objected that he could not *just* be aware of the primary qualities of the cube for, in order to be aware of the position of his limbs, he must be having some kinaesthetic sensations. The role of kinaesthetic sensations in our awareness of the position of our limbs is notoriously a matter of dispute but I would contend that knowledge of the position of our limbs is not mediated by awareness of organic sensation. For people usually know where their limbs are but are normally only conscious of their kinaesthetic sensations in certain circumstances, e.g. when they are in some discomfort or are paying special attention to what it feels like when, say, one has one's fingers crossed. However, the point could be conceded without damage to my main contention. For we are required to imagine, not a man who has no organic feelings at all, but a man who is not aware of such secondary qualities of the cube as felt temperature, felt softness or hardness, and that point remains unaffected by the concession.

McGinn has made the classic error of concentrating on sight (tactile experience is, as we have seen, just touched on at p. 85) and then generalizing his conclusions to include the other senses. However, if it is plausible to suppose that IT is true of sight but not of the other senses this disparity still stands in need of explanation. I offer the following tentative account. Visual experience is inherently spatial in a way in which auditory, olfactory and gustatory experience is not. Because it is inherently spatial it must be as of (at least) the primary qualities of extension, size and shape. Since experiences in these other three modalities are not inherently spatial it is possible for such experiences not to be as any of primary quality, for such experiences need not be experiences as of objects at all. Where an experience is of objects possessing primary qualities, such as size, shape and solidity, there must be something in the experience which enables the observer to differentiate objects, to trace their surfaces and boundaries. In the case of tactile experience I suggested this might be done solely by means of the observer knowing the position of his limbs. Nothing comparable seems available in the case of sight, which is not a sense requiring bodily contact between observer and object. It is tempting to suggest that the only mode in which we can make those discriminations in our visual field that are constitutive of awareness of primary qualities is colour perception (or something akin to it).

There could be other modes of perceptual experience which shared with visual experience the properties of being inherently spatial and not mediated by bodily contact between observer and object. Of any such experiences IT would also hold. It may be that the 'radar' sense field of a bat has these properties, although we have no conception of the nature of the secondary qualities by which the bat's primary quality experience is mediated (see Thomas Nagel, 'What is it like to be a bat?' in *Mortal Questions*, Cambridge 1979, pp. 165-80).[1]

University of Keele,
Keele, Staffordshire ST5 5BG

© DAVID MCNAUGHTON 1984

---

[1] I have benefited from discussion of these issues with Richard Swinburne.

A NOTE ON SMITH ON ATTEMPTS AND INTERNAL EVENTS

By MAURICE RICKARD

M. SMITH (ANALYSIS 43.3, June 1983) endeavours to undermine the view that all tryings are internal events (and indirectly that basic actions are internal) by reviewing certain assumptions about cross-world identity occurring in the 'paralysis' arguments commonly used to support that view. Whatever the merit of his brief comments about identity, his initial assumption, that all arguments that tryings are internal have the common fallacious form he presents, seems unconvincing. There is a slightly different route to the conclusion in question which, when developed, may well seal the gap Smith exploits in the paralysis arguments.

The view that all tryings are internal arises, says Smith, with the belief that though I may actually try to raise my arm and have my arm rise, it might have been that I was paralysed and unable to raise it, despite the very same particular attempt. Though my arm would not have risen, the trying in the possible case would have been the same event as that in the actual, and since the possible attempt would have to be inner, so too would the actual attempt. To this Smith responds that we can concede that the actual and possible attempts are the same but not that, as a consequence, the arm rising is not part of the actual successful attempt. He enlarges on this by presenting two alternatives. Assuming the actual and possible attempts to be the same is consistent with supposing that all parts of attempts are essential or else that attempts have non-essential parts. If the unsuccessful attempt is internal, then only on the former supposition will it be acceptable to claim that all attempts are internal, for then, arm rising would not be part of the attempt. On the latter, we could suggest, as Smith does, that arm rising is a non-essential part of the successful attempt, and so deny that all attempts would be totally internal. Which is the case? Without venturing into the complex area of essentialism and modal properties, it suffices here to determine just whether arm rising is part of the successful attempt. Given the above hypotheses (i.e. that the actual and possible attempts are identical and the possible is internal), if the rising is a part then the essentialist view can be rejected and Smith's adopted. Its not being a part will be grounds for discounting Smith's suggestion.

What then, would a viable individuation of successful attempts to ϕ say on the matter? Or, to put this more helpfully, what events does the language used to report attempts locate as the successful ones? First, it should be noted that if a uniform account is to emerge, successful attempts to ϕ should not be individuated differently depending on whether 'ϕ' is a basic or non-basic action description. So, whatever appears relevantly true of 'non-basic' successful attempts might then be thought true of 'basic' ones like attempted arm raisings. Further, wherever there is a difference in this respect it must be explained. To begin with a simple case, consider Roger's attempt (successful as it turns out) to raise a rather heavy flag. Before the flag begins to rise, we are likely to count his tugging the ropes as being his attempt and, more importantly, after the flag begins to rise, or has risen, we would be disinclined to revise or add to this previous judgement. Though we originally did not know the attempt would be a successful one, the rising of the flag is seen as a causal consequence of the attempt, the event for which the attempt was required rather than a late addition to it. Similarly, it would be odd to count, as part of Manuel's attempt to sink the *Sheffield*, the submerging of the ship several minutes later. If it is a part, are the intervening events leading to the sinking (i.e. the missile flight, the burning of the deck, etc.) included also? If so, Manuel appears to become busier by the second, and if not, we are left with a compound event part of which occurs with the pressing of a button and part of which mysteriously 'pops up' miles and minutes later. The more the 'goal' event is removed from the beginning of the attempt, the less credible the view that it is part of the attempt. (For those who think this is a problem only for 'non-basic' attempts, legitimate 'basic' cases can be contrived where there is delay and causally intervening events between, e.g. the beginning of an attempt to raise one's arm and the arm's rising.)

Once again, if Smith's suggestion was plausible, one would expect this to be evidenced in, or at least consistent with, the language reporting attempts, particularly in the truth conditions for such reports. Recalling Manuel's efforts, after he has pressed the missile button but before the ship eventually sinks, it seems plainly true and acceptable to say (in the past tense) 'Manuel has tried to sink the *Sheffield*'. We do not say, with it in mind that the ship's sinking would be part of his attempt if successful, that he has tried if unsuccessful but has not yet if he is successful. With respect to their parts and the temporal extension due thereto, successful and unsuccessful attempts are treated equally by the language. Proposing otherwise would require some explanation of why utterances that would thus become counterintuitively false should appear intuitively true. A view more respectful of the linguistic data would be that Manuel's pressing of the button and the ship's sinking are not each parts of his successful attempt, but are related as cause to effect, the successful attempt being whatever Manuel does that causes the ship to sink. Similarly, Roger's successful attempt would consist only in whatever he does in raising the flag, i.e. causing the flag to rise.

Heeding the earlier demand for uniformity in the individuation of 'basic' and 'non-basic' attempts, and seeing that suitable 'delay' cases could be constructed for 'basic' attempts, it behoves us to consider one's successful attempt to raise one's arm as being an event that causes the arm to rise, and no more; that is, the raising of one's arm, whatever internal event that may be.

LaTrobe University,
Bundoora, Victoria, Australia 3083

© MAURICE RICKARD 1984

ANOMALOUS MONISM AND EPIPHENOMENALISM: A REPLY TO HONDERICH

By PETER SMITH

THE fate of Davidson's version of the identity theory is an important enough issue for it to be worth commenting on Ted Honderich's brisk note 'Anomalous Monism: Reply to Smith' (ANALYSIS 43.3, June 1983; for the preceding debate see Ted Honderich, 'The Argument for Anomalous Monism', ANALYSIS 42.1, January 1982 and Peter Smith 'Bad News for Anomalous Monism?', ANALYSIS 42.4, October 1982).

The key argument of Honderich's original paper, as I read it, started from the premise (A) that the anomalous monist is committed to saying that it is, in a phrase, 'the mental as physical' which causes a given behavioural upshot. In other words, it is in virtue of the event's physical properties that a mental event causes a given action. So, since our monist maintains that (B) there are no nomological links between a thing's mental and physical characteristics, it follows that (C) there are no nomological links between a mental event's mental properties and its causal physical properties in virtue of which it produces a certain upshot. So, on the monist's view, it seems to follow that (D) it is a mere accident that an event which has the mental characteristic of being a desire, for example, should have the causal power to bring about the action which commonsense psychology would say the desire produced. And this conclusion, I would agree, is intolerable, and runs counter to our ordinary presumptions about the efficacy of the mental.

In my reply, I accepted (A) and (B) for the sake of argument, and attempted to block this reductio of the Davidsonian position after stage (C). That is, I argued that our monist can insist that there is a good sense in which it is not an *accident* that a particular mental event, identified as such, has the physical causal powers it does have, without being committed to reinstating the unwanted psycho-physical laws. Honderich, however, finds my labours to be beside the point:

> Having assumed that it is the mental event as physical which is causal, Smith notes correctly that on Davidson's view it is not a matter of lawlike connection that the event as physical *is* the event as mental. That is all that I meant by saying that it is a matter of accident that the physical event is the mental event... The proposition that it is a matter of accident is not, as Smith supposes, an independent proposition got by inference from the proposition denying lawlike connection. To say the same thing again, it is nomologically inessential, to the event's being the physical event it was, that it was the mental event it was. Hence, it is inessential to the relevant effect, the action, that the event was the mental event it was. Here we have epiphenomenalism, the inefficacy of the mental.

But this is not entirely perspicuous. Let's start by considering the accusation of epiphenomenalism. Traditionally understood, this is the doctrine (E) that mental events do not cause physical events at all. And we should immediately note that this is *incompatible* with (C) and (D). For of course, it is one thing to assert that the causal powers of mental events are not nomologically related to (or are only accidentally related to) the mental properties of those events: it is quite another thing to deny that mental events have any causal powers at all. So which does Honderich now want to say that the anomalist monist is committed to: (C/D) or (E)?

It would seem that the Davidsonian monist who identifies each mental event with some physical event must be thoroughly wedded to the assertion that mental events (at least in virtue of their physical properties) *can* and *do* cause physical events. So it is extremely difficult to see how an accusation of epiphenomenalism against the Davidsonian could possibly arise. Certainly, Honderich's argument quoted above does not sustain such an accusation. For it obviously cannot follow from the thought that 'it is inessential' to a given physical effect that its physical cause 'was the mental event it was' that the mental event in question lacks all causal powers – for by hypothesis the mental event is identical to a physical event *with* such powers. Any suggestion of epiphenomenalism here would seem to be simply confused.

It might be replied to all this that I am just being horribly pedantic and, even if 'epiphenomenalism' is traditionally used for the doctrine expressed by (E), Honderich is using the term in a permissibly stretched sense to cover the distinct but related doctrines (C) or (D); it is the suggestion that the anomalous monist is committed to one or both of these which is at the heart of his argument. But consider the following remark which Honderich makes in the course of discussing my attempt to block the argument I sketched at the outset. He argues that the monist, even if he adopts the manoeuvre I offered him,

> will presumably persist in the denial of lawlike connection noted above. It therefore remains as inessential as ever to a certain physical event's being as it is, and having the causal connections it does, that it is a mental event to which ordinary belief – but not the Anomalous Monist – assigns the same causal connections. *That* is the objection.

Here, the interpolated remark implies that the monist does *not* assign causal connections to the mental event which ordinary belief sees as causing a certain upshot. So, the monist apparently *is* being taken to be an epiphenomalist in the straightforward traditional sense. But, to repeat, this is extraordinary: the anomalous monist *of course* assigns the same causal connections to a mental event and the physical event which it is identical with: to do otherwise would be to abandon the token identity theory which is distinctively his.

Suppose, however, we ignore that troublesome interpolated remark and take it that Honderich's real complaint against the monist is *not* that he is an epiphenomalist (in the traditional sense): what we are left with is the objection that 'it is inessential' to a given physical event's causing a particular action that 'it is the mental event to which ordinary belief... assigns the same causal connections'. And, as we saw from the first quotation above, Honderich now insists quite explicitly that all he means to signify by 'inessential' here is that there is a lack of strict nomological connections. So, to make things absolutely clear, Honderich's complaint is that anomalous monism entails (C) there are no strict nomological connections between the causal physical properties of a particular mental event (in virtue of which it causes a certain action) and its mental properties (in virtue of which it is the desire, or whatever, which ordinary belief takes as the cause of the action). Well, let us grant that the Davidsonian position does indeed entail (C): *why is that supposed to be an objection?* For (C) just isn't the sort of proposition with respect to which we have clear pre-theoretic intuitions: in other words, we can't say *a priori* that (C) is obviously absurd and that any theory of the mind which entails it is thereby ruled out. For while we certainly do want there to be *some* kind of connection between an event's mental properties and its causal powers, the technical issue whether the connection in question is a strict *nomological* one is just not one that can be decided in advance of our philosophical theorizing. Compare: we certainly want there to be some kind of connection between an event's being a hurricane and its having the physical properties in virtue of which it destroys buildings – but is it clear that this connection must be nomological? (Cf. my 'Hess on reasons and causes', ANALYSIS 41.4, October 1981, the main point of which is in fact untouched by Denise Meyerson Taylor's discussion 'Actions, reasons and causal explanation', ANALYSIS 42.4, October 1982.)

Of course, if we could infer from (C) that it is a mere accident – in the ordinary, humdrum sense – that an event's mental properties are co-instantiated with its causal powers, that would settle the issue. In other words, if we could show that anomalous monism does entail (D), that would indeed be damaging. But it was precisely the point of my earlier reply to argue that we can block this move from (C) to (D). Honderich does not show that my blocking manoeuvre fails, but only insists that it is unnecessary because the damage is already done once the argument gets to stage (C). I can only reiterate that, once (C) is distinguished from (D), there is nothing particularly counterintuitive in (C) as it stands, and the mere fact that a theory entails it is not an impressive objection.

In summary: anomalous monism does entail (C), as Honderich insists, but *that* is no obstacle to accepting the position. In the light of my previous paper, I claim that monism does *not* entail the much more damaging (D). And it most certainly does not entail the epiphenomenalist's (E). I leave it to others to judge how far Honderich's reasonings depend on confusing these three things.

University College of Wales,
Hugh Owen Building,
Aberystwyth, Dyfed SY23 3DY

© PETER SMITH 1984

SMITH AND THE CHAMPION OF MAUVE

By TED HONDERICH

AS we go through the household of the Champion of Mauve, we notice many things painted mauve. Here a light-switch, there a bottle of gin, here his bedroom slippers. Take the slippers. The Champion of Mauve notes they are splendid slippers – they have the effect of keeping his feet really warm. This is the effect, he more particularly allows, of just the Hibernian fleece with which they are lined. He also allows that their being painted mauve isn't nomically necessary to their Hibernian fleece, their causal property whose effect is the keeping of his feet really warm. But, he adds, it would be a wonderful confusion to move on from that truth to any underrating of mauve in connection with his warm feet. The slippers with Hibernian fleece *are identical with* the mauve slippers. So too the light-switch with its effective circuit-breaker *is* the mauve light-switch. The bottle with the gin in it *is* the mauve bottle of gin. Let us, he adds, have no confusion about efficacy in these matters either.

Enter Smith – Peter Smith – speaking of the mind and brain. More particularly, he speaks of a man's noticing of a kipper, his wondering if it is too old, and his intending to eat it. Take the last, says Smith. In this connection, there was a neural event which was a cause of the man's then eating the kipper. More particularly, there was an event with a neural property, which property was a cause of the action. Of properties of the event, only the neural property was a cause of the action. The event did also have a mental property, that of being our man's intention to eat the kipper. That property, Smith says, was not nomically necessary to the event's having the neural property. But, says Smith, do not fall into the wonderful confusion of underrating the man's intention to eat the kipper. That intention was a property of an event that *was identical with* the event that had the neural property. Noticing and wondering about the kipper involve like facts. Do not confusedly suppose, says Smith, that what I have said offends against any conviction that we actually have about the efficacy of the mental, any conviction we actually have of the falsehood of epiphenomenalism.

My dispute with Smith, at bottom, is over whether he can distinguish himself in a relevant way from the Champion of Mauve. His 'Anomalous Monism and Epiphenomenalism: A Reply to Honderich' (ANALYSIS 44.2, March 1984, pp. 83-6) does not come near to persuading me that he can. He is committed, for all he has said, to an epiphenomalism or denial of the efficacy of the mental which is quite as impossible to accept as the mauvism of the Champion of Mauve.

He offers two speculations as to what I take epiphenomenalism to be. The first is that it is that 'mental events do not cause physical events at all'. This he understands in a way prefigured above, as follows: there are no events with mental properties, which events somehow cause physical events. But, he says, I have already accepted in setting out the argument under consideration that there are such events. His second speculation is that I take epiphenomenalism to be no more than a central proposition of Anomalous Monism as set out, that an event's mental property is not in nomological connection with its physical, causal property. This too, as he might have said, was accepted by me in setting out the argument. What he does say is that this proposition does not by itself offend against any conviction we actually have.

Speculation is all right, but reading is better. In the original piece to which Smith is objecting I wrote:

> ... there is also the other possible answer to the question of what is causally-relevant with respect to a mental event. It is the answer that the mental event as physical causes the action. To give this answer is of course to cast a new light on the first claim [of Davidson's Anomalous Monism], that the mental interracts causally with the physical. It becomes the claim that the mental as physical interracts causally with the purely physical. What is important, however, is that the resulting picture seems not to account for a conviction that lies behind acceptance of his first claim when it is naturally understood, as the claim that the mental as mental causes the physical. That is the conviction of the efficacy of the mental, already mentioned. It is the conviction that an event as mental is an ineliminable part of any full explanation of an action. It is the very root of the common denial of epiphenomenalist doctrines. ('The Argument For Anomalous Monism', ANALYSIS 42.1, January 1982, pp. 63-4)

My plain statement of what I take the conviction of the efficacy of the mental to be, and hence my understanding of the opposed doctrines, epiphenomenalisms, is repeated and considered in an article ('Nomological Dualism: Reply to Four Critics', *Inquiry*, Vol. 24, No. 3, 1981, p. 422) cited in my first rejoinder to Smith ('Anomalous Monism: Reply to Smith', ANALYSIS 42.3, June 1983, p. 149).

Smith does alight on a remark of mine in that rejoinder about 'a mental event to which ordinary belief – but not the Anomalous Monist – assigns... causal connections'. He concludes from 'but not the Anomalous Monist' that the epiphenomenalism I find in Anomalous Monism is as he first speculates, the proposition that mental events do not cause physical events at all. My words were a bit careless, but he concludes wrongly, against clear evidence the other way, above all the passage quoted above.

To repeat, Smith rightly says that in setting out a view of the mind I took on board the proposition that there are events with mental properties, which events somehow cause physical events. What I went on to do was to claim that the given view of the mind was epiphenomenalist: it did not make the mental as mental an ineliminable part of the explanation of actions. Was I there contradicting what I had earlier assumed? Obviously not. Rather, I was concluding, as I still do, that the proposition as understood that there are events with mental properties, which events somehow cause physical events, does not suffice to give us the efficacy of the mental as defined. It does not give us the mental itself as explanatory.

To look at Smith's other speculation, I do indeed suppose that there is a proposition that offends against a conviction we actually have, and that Smith's view of the mind contains that proposition. It is that mental events as mental are *not* ineliminable parts of explanations of actions. I do not suppose that we are offended by the proposition in itself that an event's mental property is not in nomological connection with its physical, causal property. We do require, of course, that something be added to that proposition to give us the efficacy of the mental.

That brings us back to the main issue. What is to be added to the given proposition to give us what we want, that the mental as mental is essential to the explanation of actions? Smith's policy here, in which he persists, is to go on saying that an event which has a mental property *is* an event which also has a physical property that is causal with respect to an action. It won't do, and no amount of repetition and emphasis will help. There are many more problems and possibilities in this neighbourhood than have surfaced in this controversy, but it is plain that the so-called identity proposition is of no more use to Smith than is the like proposition to the Champion of Mauve. The Anomalous Monist can be as wedded as he wants to the proposition that *of course* a mental event in his sense causes a physical event. By way of that truth he is no nearer getting mental efficacy than the Champion of Mauve is to getting mauvish efficacy by going on saying that it *is* the mauve slippers that keep him really warm.

Smith has another problem, which has the same source. He approaches the matter we have been considering by asking how the Anomalist Monist in general picks out the neural events with which, in his way, he then identifies mental events ('Bad News For Anomalous Monism?' ANALYSIS 42.4, October 1982, pp. 220–4). Take intendings like those of the kipper man imagined above – intendings to eat a kipper. What class of neural events does the Anomalous Monist look to? The answer given is that he looks to neural events that on his view do in virtue of their physicality actually have the causal antecedents and effects that common sense somehow assigns to the intendings, events as mental.

It is important to see what is going on here. The Anomalous Monist starts with a class of mental properties, certain intendings. He then specifies a class of physical properties, partly by their behavioural effects. Then, it is said, he 'identifies' these two classes. But of course they do not become 'identical' by fiat. More precisely, the Anomalous Monist cannot *decide* that in each instance there is one thing which has a certain mental property and a certain physical property. That is so at any rate, to be very brief, if the identity is not a mere matter of classification, as when I say the book is 328 things – the binding, some glue, and 326 pages.

Let us suppose that there really is what there must be if the view is to have a hope, some non-classificatory fact of identity or identity so-called with respect to the class of mental properties and the class of physical properties. So with noticings of kippers, wonderings about them, and so on. (The obscurity of the fact of identity is no doubt one source of the present controversy.) Why do the intendings not sometimes turn up with the kind of neural item which previously has gone with the noticings or the wonderings, or, of course, feeling melancholy about Aunt Alberta, or thinking happily of bell-jars, or feeling or thinking anything else? *What explains the non-classificatory fact of identity?* Until more is said there is a clear answer: *nothing*. This Anomalous Monism, having denied psychophysical nomic connection, is so far indistinguishable from what we all hoped had been put to rest, which is to say mere psychophysical parallelism. Will this be followed by a revival of Pre-Established Harmony? This Anomalous Monist's plight is illustrated by the fact that that would be *better*. It would offer *some* explanation of the fact in question.

I take it that Smith's defence of Anomalous Monism against my objection has not been anticipated by anyone else, as he pointed out my objection was anticipated by Frederick Stoutland. If this is so, his defence is bad news for Anomalous Monism.

University College London,
Gower Street, London WC1E 6BT

© TED HONDERICH 1984

HEMPEL REVISITED

By A. J. DALE

IN a recent paper M. L. Pokriefka ('Ayer's Definition of Empirical Significance Revisited', ANALYSIS 43.4, October 1983, pp. 166–70) has claimed – *contra* Scheffler – that it is possible to formulate a definition of *empirical significance* of 'the Ayer type' which escapes certain standard objections to which such definitions are subject. Scheffler (I. Scheffler, *The Anatomy of Inquiry*, 1963, p. 155) has contrasted definitions of empirical significance which are modifications of the original Ayer model with definitions which have an altogether different basis. It is only the former type that Scheffler believed to be incapable of eliminating, as not empirically significant, the conjunction of a significant sentence with a nonsensical one. Those conforming to the Ayer model are modifications of: S is empirically significant if and only if S, together with additional premises, logically implies some observation sentence not implied by the additional premises alone (A. J. Ayer, *Language, Truth and Logic*, Penguin, 1971, p. 15). Restrictions on the additional premises feature prominently in such modifications. In contrast, Hempel (C. G. Hempel, 'The Empiricist Criterion of Meaning', in *Logical Positivism*, ed. A. J. Ayer, p. 117, 1959) gave a very different definition of empirical significance utilizing recursion from a basis of primitive observation sentences according to the syntactic apparatus of some standard logic. In the light of these distinctions I shall argue that Pokriefka's definition is only superficially of the Ayer type and that it is in any case seriously defective.

Pokriefka's definition (M. L. Pokriefka, op. cit. p. 169) is as follows:

1. A sentence S is *directly confirmable* iff one of the following conditions is satisfied:
1.1 S is an atomic observation sentence
1.2 $S = \sim A$ and A is directly confirmable
1.3 $S = (A \& B)$, $(A \vee B)$, $(A \rightarrow B)$, or $(A \leftrightarrow B)$ and both A and B are directly confirmable
1.4 $S = (\exists x)\, F$ of $(x)\, F$ and $F(a/x)$ is directly confirmable for some individual constant a
2. S is *indirectly confirmable* iff one of the following conditions is satisfied
2.1 S is atomic and there is a directly confirmable sentence S' and a set of directly or independently established indirectly confirmable sentences Δ such that $\Delta \cup \{S\}$ is logically consistent, $\Delta \cup \{S\} \models S'$ and $\Delta \nvDash S'$
2.2–2.4 Are like 1.2–1.4 *mutatis mutandis*
3. S is *empirically significant* iff $\wedge \nvDash S$ and S is either directly or indirectly confirmable

There are a number of minor blemishes which make for inelegance: every directly confirmable sentence is indirectly confirmable, so some redundancy is inevitable; logical truths and contradictions will be directly confirmable; contradictions (though not logical truths!) will be empirically significant. With obvious modifications, however, the definition could be invisibly mended to take account of each of these.

In contrast, there is a much more serious defect revealed in the condition 2.1 under the standard positivist assumptions that atomic sentences have the form Fa and that Fa is an atomic observation sentence if and only if Fb is an atomic observation sentence where a and b are individual constants. The defect inherent in 2.1 is revealed if we attempt to use it to determine whether a given atomic non-observation sentence Fa is indirectly confirmable. Fa is indirectly confirmable if and only if a sentence S' and a set of sentences Δ satisfying the conditions of 2.1 exist. Now S' will not contain any occurrence of the predicate F since S' is *directly* confirmable and the recursive clauses of 1.2–1.4 ensure that this is so only if each atomic sentence in S' is an observation sentence whereas we are testing a non-observation sentence Fa. Suppose no member of Δ contains the predicate F. Since $\Delta \nvDash S'$ there is an interpretation of the predicates in Δ and S' such that each sentence of Δ is true and S' is false for that interpretation. Since neither Δ or S' contains F this same interpretation suffices to show that each of the sentences of $\Delta \cup \{Fa\}$ will be true under this interpretation when augmented by truth assigned to Fa. Thus $\Delta \cup \{Fa\} \nvDash S'$ which contradicts the conditions of 2.1. So for any Fa which satisfies 2.1 F must occur in at least one of the sentences of Δ. To determine in turn if such a sentence of Δ is indirectly confirmable it is necessary to determine if, for some b, Fb is indirectly confirmable since the recursion clauses 2.2–2.4 entail that any sentence containing F is indirectly confirmable only if Fb is indirectly confirmable. Thus we have come full circle. To determine whether an atomic non-observation sentence is indirectly confirmable by 2.1 we would first have to determine if another atomic non-observation sentence containing the same predicate F is indirectly confirmable! Such a procedure would never end. As a test for confirmability clause 2 of Pokriefka's definition is totally useless. It would seem that the definition collapses into its first clause 1–1.4 which is simply a recursion from primitive observation sentences. Such a definition is almost identical to that given by Hempel (C. G. Hempel, op. cit. p. 117) – Hempel's was given in terms of predicates rather than sentences and included, in addition to those syntactic resources mentioned in the above definition, the construction 'the class of *Fs*'. As such a recursive definition it has the disadvantages as well as the advantages of Hempel's original: it eliminates the possibility of significance for the conjunction of a significant sentence with a nonsensical one only at the cost of severely restricting what is to count as empirically significant – dispositional properties, for example, will produce grave difficulties for such accounts. Be that as it may, there is no doubt that Pokriefka's definition, once shorn of the clauses 2–2.4, as I have shown it should be, will no longer be of the Ayer type since the only similarity it possesses with such definitions is to be found in 2.1 of that definition.

The University, Hull HU6 7RX © A. J. DALE 1984

MORE ON EMPIRICAL SIGNIFICANCE

By M. L. POKRIEFKA

IN [1], I amended A. J. Ayer's definition of empirical significance to meet the criticism that it confers significance on any sentence whatsoever. That my definition disposes of the targeted criticism can be rigorously proved. What I failed to notice is that the definition is intolerably restrictive. It banishes metaphysics from the realm of empirical significance only at the expense of taking most science along for the ride. In the present note I show that this is so.

Consider a formal language constructed in the customary way from quantifiers, truth functional connectives, n-ary predicate letters for arbitrary n, individual constants, and parentheses. A definition of empirical significance can be cooked up for this language according to the recipe given in [1]. That definition will be too restrictive if it can be shown that

Thm. If a sentence S contains at least one occurrence of a predicate G that is not an observation predicate, then S is not empirically significant.

Proof. It is obvious that S is empirically significant only if G occurs in some indirectly confirmable atomic sentence T. Consider, now, any nonempty set of directly or independently established indirectly confirmable sentences Δ such that $\Delta \cup \{T\}$ is logically consistent. Since no atomic sentence containing an occurrence of G has been independently established as indirectly confirmable, no $P \in \Delta$ contains an occurrence of G. So, for each S' such that $\Delta \cup \{T\} \nvDash S'$ and $\Delta \nvDash S'$, G must occur in S'. Hence, S' is not directly confirmable. Therefore, T is not indirectly confirmable.

The theorem shows that only sentences all of whose predicates are observation predicates can satisfy my definition. Thus, the definition does demarcate a class of empirically significant sentences, but the members of that class are a 'fancifully fancyless medium of unvarnished news' indeed.[1]

Michigan State University,
East Lansing, Michigan 48824, U.S.A.

© M. L. POKRIEFKA 1984

REFERENCES

[1] M. L. Pokriefka, 'Ayer's Definition of Empirical Significance Revisited', *Analysis*, 43.4, October 1983, pp. 166-70.

---

[1] I would like to thank Herbert Hendry for helpful discussion.

THE TIGER AND HIS STRIPES

By WILLIAM LYONS

IN a passage in *Content and Consciousness* (Routledge, 1969), Daniel Dennett put forward an argument against the mental image theory of imagination in the following terms:

> Consider the Tiger and his Stripes. I can dream, imagine or see a striped tiger, but must the tiger I experience have a particular number of stripes? If seeing or imagining is having a mental image, then the image of the tiger *must* – obeying the rules of images in general – reveal a definite number of stripes showing, and one should be able to pin this down with such questions as 'more than ten?', 'less than twenty?'... Of course in the case of actually seeing a tiger, it will often be possible to corner the tiger and count his stripes, but then one is counting real tiger stripes, not stripes on a mental image. (pp. 136-37)

The implication, of course, is that one cannot count the stripes on an imagined tiger, so it cannot be the case that imagination involves images.

Now the usual reply to this argument of Dennett is that one cannot count the stripes on an imagined tiger, not because an imagined tiger is not really an image, but because it is a poor, indistinct image. Thus, for example, Alastair Hannay suggests that counting the stripes is impossible because an imagined tiger is 'fluttering, febrile, vanishing', 'evanescent', 'vague' and 'fleeting' (*Mental Images – A Defence*, Allen & Unwin, 1971, Ch. V, Section 3) while J. A. Fodor argues for the same conclusion because an imagined tiger is 'labile', 'blurred', and there are 'many visual properties which would not be pictured' (*The Language of Thought*, Crowell, 1975, Ch. 4, pp. 187-91).

I want to suggest a better defence against Dennett's argument, for I shall attempt to show that both Dennett's argument and the usual reply to it miss a fundamental point about the nature of imagination. Put bluntly, my suggestion is that one cannot be more sophisticated visually in imagination than one is in perception itself because imagination is parasitic on perception, and that if this principle is attended to, one will realise that one *can* count stripes on imagined tigers.

But let us go back to the beginning. Instead of cornering a tiger (which would *not* be the time to attempt to count his stripes!) let us settle him down to his first hefty feed of the day. Now let each of us move closer and attend to the task of counting his stripes (given that we have decided on a procedure for individuating stripes). Let us suppose also that the tiger in fact has 273 stripes. Now one cannot see, much less count, 273 stripes in one go. So one scans slowly up the hind legs, along the sides, and so on. The bravest among us might stick out a finger to help the process along. The point is, in ordinary perception, counting stripes will involve a large number of focused perceptions; 273 if we take each stripe singly (less if we feel we can take them in twos, threes, etc.).

Now, in regard to imagination, we can expect no better than what is possible in perception, for it is sensible to suppose that, whatever be the pictorial or image-like nature of imagination, it is pictorial or image-like in so far as it is based in some way as yet unknown upon ordinary perception. So to count stripes in imagination will involve a series of images of single stripes (perhaps with our finger pictured as picking out the stripes), repeated 273 times. That is how to count stripes in imagination because that is the only way we can count them in ordinary perception.

Dennett, Hannay and Fodor all seem to presume that, if imagining a tiger involves an image of a tiger, then counting the stripes on an imagined tiger must involve the fruitless task of holding that image very still, and then trying steadily to pick out stripes on the side of this tiger-image, and finally trying to count them. This is very revealing. In perception, there is a holding steady and a working over when counting stripes on a tiger but *not* a holding steady of a perceptual 'snapshot' of the whole tiger over which we then work steadily. What we hold still in order to work over with a long series of focuses on particular stripes will be a real whole tiger which we have steadied in some way, we hope. Since imagination is parasitic on perception, then the imagined counting of stripes can at most be a repeat in imagination of what we do in perception, that is, engage in a long series of 'snapshots' of single, fingered stripes.

We do not have a real tiger in our heads to hold steady, and any single image of an imagined tiger will be of no more use for counting than would a single holistic percept of a tiger. In a single holistic percept, we can only take in, at most, a few salient tiger features – that it is striped (but not how many stripes there are, nor the precise hues), that it is quiet and eating (but not exactly what it is eating), etc. – so we would get nowhere at all if, to count stripes on a real tiger, we had to use our few-salient-features, 'snapshot' percept of the whole tiger rather than the real tiger. Likewise we will get nowhere at all if, in imagination, we try to count stripes by concentrating on an image of a whole tiger.

To demand of imagination that the one imagining be able to hold steady the 'picture', which is an imagined tiger, and then run a finger along the body of the pictured tiger, and thereby to count the stripes, is to misunderstand not merely the nature of imagination but also the nature of perception. One can count stripes on imagined tigers but one must do so in the same way as one counts stripes on real tigers.

The University, Glasgow G12 8QQ © WILLIAM LYONS 1984

GREEN ON DICTATORS AND DEMOCRACIES

By DAVID GORDON

LESLIE GREEN, in 'Dictators and Democracies' (ANALYSIS 43.1, January 1983) claims that one of the two following precepts should be abandoned: '(1) a social choice rule is democratic only if there exists no dictator' or 'the economically rational individual should (2) participate in a social choice only if the benefits of participation outweigh the costs' (p. 58). Green argues that, by (2), a rational voter ought to vote only if the benefits of participation in voting, multiplied by the probability of his vote swinging the outcome, exceed the costs to him of voting. Such a calculation will result in its being rational to vote only in very small electorates, since in 'any large electorate p (the probability of one's vote swinging the outcome) approaches zero' (p. 59).

One may, according to Green, go further. The benefits of voting, multiplied by p, will be likely to exceed the costs only in electorates of one. 'But if there is an electorate of one then the efficacy of the preferences of those individuals who have no vote is zero. If one individual's preferences have a weight of one and all other preferences weigh zero, then there exists an individual whose preferences are automatically society's preferences independent of the preferences of all other individuals, and therefore the condition of non-dictatorship is not satisfied. Thus the rational voter should (3) participate in social choice only if there is a dictator' (p. 59). From (1) and (3), furthermore, it follows that 'the rational voter should (4) participate in a social choice only if it is not democratic' (p. 59).

If it were objected that a benefit cost calculation will show it is rational to vote in more cases than that of an electorate of one, Green answers that, assuming the correctness of this point, a weaker version of his argument still holds. '(3) becomes (3*): there is *best reason* to participate in a social choice when there is a dictator. (4) would be similarly weakened but not, I think, much less embarrassing' (p. 59).

It seems to me that Green's argument depends on an overly broad notion of what situations are dictatorial. Suppose that there is a club of ten members. The rules of the club allow the president to nominate the vice-president, regardless of the preferences of the remaining members. Here it would be right to say that the president is a dictator for the club so far as the choice of vice-president is concerned. But suppose instead that according to the rules of the club, the vice-president is to be chosen by a majority vote of the members present at a specified meeting. At the electoral meeting, only one member appears and his vote is accordingly decisive. (The other members are not prevented from attending: they simply choose not to come.)

Here, although it is true the efficacy of the preferences of the absent members is zero, I do not think it right to say that the one member who voted is a dictator. The others *could have* voted: it was not specified in the club rules that the preferences of the particular individual who voted were to be decisive. In point of fact, the rules specified the contrary: any member who was present at the specified meeting would have had a vote. Given these rules, it is surely counterintuitive to say that the voter is a dictator. If this is correct, then (3) should be reformulated to

(3′) a rational voter should participate in a social choice only if he alone is making the choice.

Since, in most elections where more than one person is eligible to vote, the probability of only one voter voting will be extremely small, it follows that, if (3′) is correct, it will rarely, if ever, be rational to vote. But this result does not require that (1) be abandoned since the condition in (3′) can be true without the existence of a dictatorship. There need be no individual 'whose preferences are *automatically* society's preferences independent of the preferences of all other individuals' (p. 59, emphasis added) for the condition in (3′) to hold.

One may see the fact that Green has failed to show an inconsistency between (1) and (3′) even more clearly if one considers the weaker version of his argument. If my point about dictatorship is right, (3*) becomes (3*′), 'there is *best reason* for someone to participate in a social choice if he thinks his preferences will be absolutely decisive'. And there is nothing whatever paradoxical about this: (3*′) is an obvious truth.

6238 Orange Street,
Los Angeles, California 90048, U.S.A.

© DAVID GORDON 1984

Philosophical Investigations

Edited by **D.Z. Phillips**

Philosophical Investigations publishes articles which discuss, extend or criticise philosophical approaches pioneered by J. L. Austin, Gilbert Ryle and Ludwig Wittgenstein.

Volume 7 No. 2 April 1984

Articles:
The Language of Sense Data and Private Experience – II
Notes taken by RUSH RHEES of Wittgenstein's Lectures, 1936
The Logic of Intentional Verbs DAVID CARR

Critical Notice:
John H. Whittaker, *Matters of Faith and Matters of Principle* WALFORD GEALY

Reviews:
R. M. Hare, *Moral Thinking* COLIN LYAS
Nathan U. Salmon, *Reference and Essence* MICHAEL COHEN
Philip Bean, *Punishment* ANTHONY MANSER
Quentin Lauer, *Hegel's Concept of God*
Ronald Bruzina and Bruce Wilshire (eds.) *Phenomenology: dialogues and bridges* DAVID LAMB

Published quarterly: January, April, July and October
Subscriptions Volume 7 (1984)
Institutions: *£25.00 (UK); £30.75 (Overseas); $58.50 (US); $71.70 (Canada)*
Individuals: *£11.95 (UK); £14.65 (Overseas); $28.00 (US); $34.30 (Canada)*
Single issues are available; prices on request.

Orders, with remittance, or requests for specimen copies and further details should be sent to:
Sue Dommett, Journals Department, Basil Blackwell,
108 Cowley Road, Oxford, OX4 1JF, England.

Basil Blackwell · Oxford · England

The Philosophical Quarterly

Vol.34 No.135 APRIL 1984
© The Philosophical Quarterly 1984

Edited in the University of St. Andrews by an Editorial Board
Chairman of the Board of Editors: LESLIE STEVENSON
Executive Editor: GORDON GRAHAM *Reviews Editor*: ANTHONY ELLIS

Anti-Realism Unrealized — *Anthony Appiah*
Aristotle's Now — *Sarah Waterlow*
Hume's Two Standards of Taste — *Jeffrey Wieand*

DISCUSSIONS
Popper's Solution of the Problem of Induction — *Colin Howson*
Jackson on Physical Information and Qualia — *Terence Horgan*
The Disjunctive Syllogism and Subjunctive Conditionals — *A.J. Dale*
Reply to Priest on Being's Paradox — *Ross T. Brady*
Sensible Self Containment — *B.H. Slater*

BOOK REVIEWS

Subscriptions for 1984

Annual Subscription	UK	Overseas	USA	Canada
Institutions	£22.50	£27.50	$50.00	$62.00
Individuals	£10.25	£12.50	$25.00	$30.00

Orders, remittances and enquiries should be addressed to: Journals Department,
Basil Blackwell Publisher Ltd., 108 Cowley Road, Oxford OX4 1JF, England.

THEORIA

A Swedish journal of philosophy
Edited by Peter Gärdenfors, Bengt Hansson
and Krister Segerberg

Volume 49, part 1
SPECIAL ISSUE ON EVOLUTIONARY EPISTEMOLOGY

A. Olding Biology and knowledge
Andrew J. Clark Meaning and evolutionary epistemology
Neil Tennant In defence of evolutionary epistemology

THEORIA, which invites contributions in all fields of analytic philosophy, usually publishes three issues a year. Subscriptions and requests for back numbers should be sent to *Editor,* THEORIA, *Filosofiska institutionen, Kungshuset i Lundagård, S-223 50 Lund, Sweden.* Subscription for one year is SwCr 60:– or U.S. $ 16.–, for three years SwCr 160:– or U.S. $ 40.–.

SOCIAL PHILOSOPHY & POLICY

NEW

Editor: Ellen Frankel Paul
Associate Editors: Fred D. Miller and Jeffrey Paul

Volume 1, no. 2 March 1984

Human Rights

ALAN GEWIRTH	The Epistemology of Human Rights
ARTHUR C. DANTO	Constructing an Epistemology of Human Rights: A Pseudo Problem?
ALAN GEWIRTH	A Reply to Danto
LOREN LOMASKY	Personal Projects as the Foundation for Basic Rights
HUGO ADAM BEDAU	Why do we Have the Rights We Do?
JOHN GRAY	Indirect Utility and Fundamental Rights
ALLAN GIBBARD	Utilitarianism and Human Rights
JAMES FISHKIN	Utilitarianism Versus Human Rights
DAVID KELLEY	Life, Liberty, and Property
MARTIN P. GOLDING	The Primacy of Welfare Rights
ALAN DONAGAN	The Right Not to Incriminate Oneself
RICHARD E. FLATHMAN	Moderating Rights
CHARLES R. BEITZ	Difficulties with Flathman's Moderation Thesis

Subscription Rates
Volume 1: Autumn 1983 and Spring 1984
Individuals £10.50 (UK) £12.00 (overseas inc. Eire)
$14.50 (USA) $17.75 (Canada)
Institutions £24.50 (UK) £29.50 (overseas inc. Eire)
$40.00 (USA) $49.00 (Canada)

Orders, with payment, or sample copy requests to: **Sue Dommett,**
Basil Blackwell, 108 Cowley Road, Oxford, OX4 1JF.

Basil Blackwell for the
Social Philosophy & Policy Center

Philosophical Books

***Philosophical Books** aims to provide prompt, scholarly reviews of professional books in philosophy. Besides conventional reviews in both contemporary philosophy and the history of philosophy, it now includes regular Discussion features, where the authors of selected titles have the opportunity to reply to reviewers. There are occasional subject surveys and issues devoted to a single theme.*

Books for review should be sent to the Editor, Dr M.A. Stewart, Department of Philosophy, University of Lancaster, Lancaster LA1 4YT.

Published quarterly: January, April, July and October
Subscriptions to Volume 25 (1984)

Institutions: *£21.00 (UK) £25.75 (overseas) $52.50 (US) $64.30 (Canada)*
Individuals: *£10.95 (UK) £13.60 (overseas) $26.95 (US) $33.00 (Canada)*
Single issues: *£5.80 (UK) £7.20 (overseas) $13.50 (US) $16.75 (Canada)*

Orders, with remittance, or requests for specimen copies and further details should be sent to: Sue Dommett, Journals Department, Basil Blackwell, 108 Cowley Road, Oxford, OX4 1JF, England.

BASIL BLACKWELL · OXFORD

Journal for the Theory of Social Behaviour

General Editors: C.W. Smith, P. Collett,
and S. Moscovici

Volume 14 no. 1 March 1984

Contents
Relativism, History and Objectivity in Human Studies JOSEPH MARGOLIS
Domains of Applicability of Social-Scientific Theories: Problems in the Empirical Falsifiability of Bounded Generalizations PETER KNAPP
The Psychology of Abstraction DAVID KELLY and JANET KRUEGER
The Kantian Revolution in Perception AARON BEN ZEEV
Mutual Intention RICHARD POWER
Analytic Work: Aspects of the Organisation of Conversational Data R.J. ANDERSON and W.W. SHARROCK
B.F. Skinner's Accounts of Private Events: A Critique ROBERT G. BURTON

1984 subscriptions:
Institutions £28.95 (inland) £35.50 (overseas) $74.00 (USA) $90.60 (Canada);
Individuals £12.95 (inland) £15.95 (overseas) $35.95 (USA) $44.00 (Canada)
Single copies are available, please write for prices.

Orders to: JOURNALS DEPARTMENT, **BASIL BLACKWELL,**
108 Cowley Road, Oxford OX4 1JF, England.

Philosophy

Editor: RENFORD BAMBROUGH *University of Cambridge*
Published for the Royal Institute of Philosophy

The journal is concerned with the study of philosophy in all its branches. Two supplements: *Objectivity and Cultural Divergents* and *Philosophy in Practice* will be published during 1984 and are included in the subscription price.

Recent articles include: SOPHIE BOTROS: Acceptance and Morality; JOHN COTTINGHAM: Neonaturalism and its pitfalls; MARVIN GLASS: Not going to Hell on one's own; RALPH W. CLARK: Induction justified (but just barely); CATHERINE WILSON: Literature and knowledge; A. PHILLIPS GRIFFITHS; Hare's *Moral Thinking*

Volume 59: January, April, July and October 1984
£48.00 Single parts £9.00 Airmail £13.50 extra
ISSN 0031-8205

For further information, please write to Aileen Askwith, Journals Publicity Controller at the Press.

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Shaftesbury Road, Cambridge CB2 2RU, England

MIND

A QUARTERLY REVIEW OF PHILOSOPHY
edited by D. W. Hamlyn
with the assistance of S. D. Guttenplan

Vol XCIII No 370 April 1984

Articles

Seeing the Present: J. BUTTERFIELD
Rigidity and Scope: A.D. SMITH
On the Nature of Rights: J. RAZ
Events and Explicative Definitions: G.N. SCHLESINGER
An Evidentiary Theory of Promises: J.P.W. CARTWRIGHT

Discussions

The Principle of Equality: A Reply to Nathan: T. HONDERICH
Lowe on Modalities *De Re*: T. BALDWIN
Frege on Thoughts: A Reply: G. CURRIE
Why Nozick is a Sceptic: A.L. BRUECKNER
Why Time is Extensive: G. PLUMER
Consciousness and the Exclusivity of Function: E.R. KRAEMER

Critical Notice

The Interpretation of Frege's Philosophy by Michael Dummett: D.R. BELL

Book Reviews Books Received Notices

Orders to: Journals Department,
Basil Blackwell,
108 Cowley Road,
Oxford OX4 1JF, England

Annual subscription:
Institutions: £14.50 or $32.50 (Canada $37.75)
Individuals: £10.00 or $22.50 (Canada $26.25)
Single issues: £4.00 or $9.00 (Canada $10.50)

Published for the MIND ASSOCIATION by BASIL BLACKWELL

INQUIRY *Editor: Alastair Hannay*

Vol. 26 (1983), No. 3

Noumenalism and Einstein's Argument for the Existence of God	Lewis S. Feuer
Marx and Cohen on Exploitation and the Labor Theory of Value	Nancy Holmstrom
More on Exploitation and the Labour Theory of Value	G.A. Cohen
The Epistemological Status of a Naturalized Epistemology	Ron Amundson
Interpreting Frege: A Reply to Michael Dummett	Gregory Currie
The Value of Wild Nature	Robert Elliot
Recent Work on Frege	Crispin Wright

Vol. 26 (1983), No. 4

Symposium on Alasdair MacIntyre's *After Virtue*	
I. Kant after Virtue	Onora O'Neill
II. Virtues, Human Good, and the Unity of a Life	Raimond Gaita
III. Morals, Moore, and MacIntyre	Stephen R.L. Clark
IV. Moral Rationality, Tradition, and Aristotle	Alasdair MacIntyre
Values in Mackie's Error Theory of Ethics	Timo Airaksinen

Subscription price:
Institutions: N.kr. 270,— $45.00 — Individuals N.kr. 140,— $23.00.

Published quarterly by Universitetsforlaget, P.O. Box 2959 Tøyen, Oslo 6, Norway
— P.O. Box 258, Irvington-on-Hudson, New York 10533, USA.

NOTES

The **ANALYSIS** Committee consists of: Chairman, P. T. Geach; Secretary, T. R. Baldwin; J. H. Benson, Margaret A. Boden, A. E. Fisher, Andrew Harrison, R. H. Holland, J. H. McDowell, D. H. Mellor, R. G. Swinburne, A. R. White, C. J. F. Williams, Peter Winch. This committee is responsible for appointing and advising the Editor and for the general policy of the paper.

SUBSCRIPTIONS. **ANALYSIS** is published by Basil Blackwell Publisher Ltd, 108 Cowley Road, Oxford OX4 1JF, England. Each volume comprises four numbers, each of at least 48 pages, appearing in January, March, June and October. Orders should be sent to the publisher. Subscription prices are as follows:

	U.K.	Overseas	U.S.A.	Canada
Institutions	£11.00	£13.50	$33.00	$39.50
Individuals	£7.00	£8.60	$21.00	$25.00

U.S. MAILING. Second class postage paid at New York. Postmaster: Send address corrections to: Analysis: c/o Expediters of the Printed Word Ltd, 527 Madison Avenue, Suite 1217, New York, NY 10022 (U.S. Mailing Agent).

CONTRIBUTIONS. Articles submitted for publication should be addressed to Mr Christopher Kirwan, Exeter College, Oxford OX1 3DP.

Contributors are asked to note the following requirements. Articles should normally not exceed 3,000 words in length, including references and footnotes. They should be type written in double spacing on one side of the paper only; photographic copies are acceptable. Footnotes should be kept to a minimum and wherever possible avoided altogether. Inset quotations and formulae should be well spaced from the text preceding and following them. The spacing of symbols within formulae should be carefully attended to. Single quotation marks should be used, except for purposes of internal quotation and "scare" quotes. Discussion papers should be sent in as soon as possible after the appearance of the article to which they refer.

Typescripts of rejected articles will be returned only if self-addressed envelopes with correct postage are provided (in the case of overseas contributors, international reply coupons sufficient for airmail).

Proofs of accepted articles will be sent to authors for correction, together with information about free copies. Typescripts of accepted articles will be retained by the Editor on the assumption that authors have kept their own copies.

Apart from correction of printers' errors, only very minor changes in proof will normally be accepted by the Editor, to avoid delay and expense in production.

29 NOV 1984

Typeset by Unicus Graphics Ltd, Horsham. Printed and bound in Great Britain by
Whitstable Litho Ltd, Whitstable

l. 44 No. 3 (New Series No. 203) June 1984

ANALYSIS

Edited by
CHRISTOPHER KIRWAN

CONTENTS

ANALYSIS competition — 'Problem' No. 20
Were those disproofs I saw before me? — NEIL TENNANT
Sainsbury on a Fregean argument — STIG ALSTRUP RASMUSSEN
Rejoinder to Rasmussen — R. M. SAINSBURY
Impossible worlds — NATHAN SALMON
Indefinite identity: a reply to Broome — H. W. NOONAN
On the tracks of the sceptic — JONATHAN DANCY
Special relativity and the location of mental events — DAVID GORDON
Reply to Gordon — MICHAEL LOCKWOOD
Preference and transitivity — RICHARD LEE
Putnam's model-theoretic argument against metaphysical realism — ANTHONY L. BRUECKNER
Phenomenalism and observation conditions — ERIC RUSSERT KRAEMER
On making others do good — M. L. POKRIEFKA
Correction

BASIL BLACKWELL PUBLISHER · OXFORD

ISSN 0003-2638

This Publication
is available in Microform.

University
Microfilms
International

Please send additional information
for ______________________
(name of publication)
Name______________________
Institution______________________
Street______________________
City______________________
State__________ Zip__________

300 North Zeeb Road
Dept. P.R.
Ann Arbor, Mi. 48106

LIBRARY

Basil Blackwell

Personal Identity

SYDNEY SHOEMAKER and RICHARD SWINBURNE

What does it mean to say that this person at this time is 'the same' as that person at an earlier time? If the brain is damaged or the memory lost, how far does a person's identity continue? In this book two eminent philosophers develop very different approaches to the problem.
Great Debates in Philosophy
192 pages, hardback **£15.00** (0 631 13208 2)
paperback **£5.50** (0 631 13432 8)

Consciousness and Causality

A Debate on the Nature of Mind

D.M. ARMSTRONG and NORMAN MALCOLM

How are the objects of consciousness related to consciousness? Can consciousness be defined as an awareness of inner mental states. Can one perceive consciousness in oneself? Can states of consciousness be identified as brain states? David Armstrong and Norman Malcolm here present opposing views on these and related questions on the philosophy of mind.
Great Debates in Philosophy
232 pages, hardback **£17.50** (0 631 13212 0)
paperback **£6.50** (0 631 13433 6)

The Legacy of Wittgenstein

ANTHONY KENNY

In these marvellously clear essays, Anthony Kenny provides an extensive commentary on Wittgenstein's work. Going beyond exegesis, he also gives a critique of other philosophers and theorists — Chomsky, Gregory, Anscombe, Teilhard de Chardin, for example — in the light of Wittgenstein's discoveries. A general introduction assesses the contemporary relevance of his thinking.
208 pages, **£17.50** (0 631 13705 X)

Wittgenstein on Meaning

COLIN McGINN

This book puts forward an original interpretation of the *Philosophical Investigations* and related works by Wittgenstein. It examines in detail what he actually wrote about meaning, understanding, rules and private language. Wittgenstein, it is argued, was opposed primarily to the 'inner process' conception of understanding, and his alternative view — that following a rule is a technique or custom — is elucidated.
160 pages, **£10.50** (0 631 13764 5)

Basil Blackwell

Property and Political Theory

ALAN RYAN

Until a few years ago, arguments about the rights of property were central to political theory. Alan Ryan brings to life the cut and thrust of these arguments from Locke and Rousseau, through Kant and Hegel, to Mill and Marx, and gives a spirited and controversial answer to the question underlying the whole discussion: why did Mill and Marx get it wrong?
240 pages, **£15.00** (0 631 13691 6)

Language, Sense and Nonsense

G.P. BAKER and P.M.S. HACKER

This is a thorough examination of the fundamental presuppositions of philosophical semantics. Inspired by Frege, Wittgenstein, Cannap and Tarski, this active branch of philosophy is now developed along lines sketched by such as Quine, Davidson and Dummett. This book offers both an historical and a conceptual analysis of the guiding ideas for a theory of meaning for a natural language.
416 pages, **£22.50** (0 631 13529 7)

But is it Art?

B.R. TILGHMAN

The question of how we identify, describe and evaluate works of art has exercised philosophers through the ages. In this lucid, witty and thought-provoking book, Professor Tilghman explores the possibility or impossibility of constructing a proper philosophical theory of art.
256 pages, **£15.00** (0 631 13663 0)

The Rise and Decline of Western Liberalism

ANTHONY ARBLASTER

This is the first major study of liberalism to appear for many years. No other book gives an overall perspective on Western liberal ideas from their beginnings in the Renaissance and Reformation to the present day, or comes to grips with the history of liberalism in the twentieth century. It is both an historical study and a critical analysis.
416 pages, **£19.50** (0 85520 765 5)

Basil Blackwell

Truth, Knowledge and Modality

Philosophical Papers, Volume 3

G.H. von WRIGHT

In this third volume, Professor von Wright concentrates on the ideas of truth, knowledge and modality, and the articles included are possibly more original and more recent than those of the earlier volumes. In subject matter, he is closer here to the topics which concern him most, and which will be of greatest interest to other philosophers.

256 pages, **£17.50** (0 631 13367 4)

From Knowledge to Wisdom

A Revolution in the Aims and Methods of Science

NICHOLAS MAXWELL

In this wide-ranging yet closely argued book, the author argues forcefully the need for a comprehensive revolution in the aims and methods of science. He urges a new kind of enquiry which would have as its fundamental aim the enhancement of personal and social wisdom.

320 pages, **£19.50** (0 631 13602 9)

Scepticism, Rules and Language

G.P. BAKER and P.M.S. HACKER

Rule-scepticism and the alleged 'sceptical solution' to its paradoxes have recently enjoyed the philosophical limelight. Here the authors argue that the problem and the solution alike are bred of confusions, misinterpretations and radical incoherences. Rule-scepticism, they suggest, is best returned to the murky backstage of philosophical nonsense whence it emerged.

160 pages, **£15.00** (0 631 13614 2)

God, Jesus and Belief

The Legacy of Theism

STEWART R. SUTHERLAND

What way forward is there for Christian faith? To reject traditional faith wholesale is to squander a precious legacy: to defend it against the light of reason displays intellectual cowardice. Professor Sutherland here combines philosophy and theology to outline a personal creed which is both radical and revisionary. It may infuriate or inspire, but it cannot fail to be thought-provoking.

224 pages, hardback **£15.00** (0 631 13548 0)
paperback **£5.95** (0 631 13591 X)

Basil Blackwell

Realism and Truth

MICHAEL DEVITT

This book argues that realism is essentially a metaphysical issue. Drawing on the correspondence theory of truth, Michael Devitt develops a strong defence of realism, which leads him to advance a naturalized approach to semantics and epistemology. In the discussion which ensues he examines and rejects in turn the views of most influential critics of realism.
304 pages, **£19.50** (0 631 13535 9)

Hannah Arendt

Politics, Conscience, Evil
GEORGE KATEB

In this exceptional intellectual biography George Kateb examines the life and work of Hannah Arendt, from the classic *Origins of Totalitarianism* (1951) to her death in 1975. A sensitive yet critical analysis, it illuminates her innovations and originality, her concentration on the idea of political action as an end in itself, her standards for political greatness, and the complex and disturbing relations she saw between political action and morality.
224 pages, **£19.50** (0 85520 691 8)

The Gramscian Challenge

Coercion and Consent in Marxist Political Theory
JOHN HOFFMAN

How far does the 'dictatorship of the proletariat' differ from the 'dull compulsion' of capitalist social relations described in *Das Kapital*? Tangible issues such as these lie at the heart of John Hoffman's analysis of coercion and consent. In a study which illuminates key problem areas in Marxist political theory and strategy and examines the problems of contemporary socialism.
240 pages, **£17.50** (0 85520 771 X)

Metaphysics

Methods and Techniques
GEORGE N. SCHLESINGER

What is metaphysics? What methods and techniques does it imply? What kinds of problem fall within its scope? This lively introduction explores these questions in terms that students can understand and at the same time brings new insights to bear on perennial philosophical debates.
278 pages, hardback **£18.50** (0 631 13124 8)
paperback **£6.95** (0 631 13125 6)

PHILOSOPHIA
PHILOSOPHICAL QUARTERLY OF ISRAEL

Volume 13, Nos. 3-4 (October 1983)

A Symposium on the Nature of Philosophy:

F. Sontag	The Program of Philosophy
J. Kekes	Philosophy, Historicism, and Foundationalism
P.T. Geach	Comparatives
E. Martin, Jr.	Fregean Incompleteness
E.M. Zemach	Identity and Open Texture
A. Margalit and M. Bar-Hillel	Expecting the Unexpected
K. Taylor	Worlds in Collision
Kai Nielsen	Formulating Egalitarianism: Animadversions on Berlin
M.J. Zimmerman	Mill and the Consistency of Hedonism
T.R. Machan	Ethics and the Regulation of Professional Ethics

Book reviews and critical studies: Felix Grayeff^{+}, Yuval Lurie, O.H. Green, Ashok Vohra, Herbert Moskowitz, F. Gunthner and Mark Vorobej.

ANALYSIS 44.3 JUNE 1984

ANALYSIS 'PROBLEM' NO. 20

The twentieth problem is set by Professor Martin Hollis of the University of East Anglia, and is as follows:

> Syme bit off another fragment of the dark-coloured bread, chewed it briefly, and went on:
>
> 'Don't you see that the whole aim of Newspeak is to narrow the range of thought? In the end we shall make thoughtcrime literally impossible, because there will be no words in which to express it. Every concept that can ever be needed, will be expressed by exactly *one* word, with its meaning rigidly defined and all its subsidiary meanings rubbed out and forgotten.' (*Nineteen Eighty-four*, part I, chapter 5)
>
> The principles of Newspeak, and especially the division of words into three distinct classes, are described by George Orwell in the appendix to the book. Does the project sketched by Syme have any philosophical snag which could not be remedied?

The word limit is 600 words. Entries should reach the Editor of ANALYSIS by 31 October 1984; they should not be sent to Professor Hollis. Entries will not be returned unless accompanied by sufficient British stamps or international postage coupons. Entrants may use their own names or pseudonyms. Entrants must be under the age of thirty, or undergraduates or graduate students.

A report with any winning entries will be published in volume 45 of ANALYSIS. The ANALYSIS Committee has voted a sum of £150 which will be awarded as a prize if Professor Hollis finds a sufficiently deserving entry.

WERE THOSE DISPROOFS I SAW BEFORE ME?

By NEIL TENNANT

HOW is the manifestation requirement in the theory of meaning seen to conflict with realism? McGinn addressed this question in 'Truth and use' in a manner that I criticized in 'Is this a proof I see before me?' McGinn defended himself in a reply whose gist is rehearsed (but not endorsed) by Weir in 'Truth conditions and truth values'. Now I believe that the question as to who is right in this matter is not verification transcendent; and in this paper I shall try to provide renewed warrant for earlier semantical claims in 'Is this a proof I see before me?'.

The manifestation requirement is as follows:

> Grasp by a speaker of the meaning of a sentence should be fully manifestable in observable exercises of recognitional capacities concerning it.

This requirement turns the theory of meaning into a theory of understanding: an account of sentence meanings is to be given in terms of what it is that a speaker knows when he understands those sentences. The mention of recognitional capacities is to alert one to the need to manifest knowledge which is implicit, rather than knowledge explicitly formulated in further sentences.

The main claim of realist semantics is that the central property of sentences with which a molecular semantic theory should be concerned is that of being (classically) true, even if a sentence's being true is recognition transcendent. Let us refer to this claim as 'Realism' and to the previous claim as 'Manifestation' for convenience. By 'Fact' I shall refer to the claim accepted by all parties to the dispute, to the effect that we understand some sentences that are undecidable – sentences, that is to say, whose truth value we do not know how to determine by any effective means at present. (We might follow Stig Andur Pedersen in calling such sentences *pro tempore* undecidable, as opposed to undecidable *tout court*.)

To Dummett belongs the credit for showing that Manifestation plus Realism plus Fact is inconsistent. In briefest outline, his argument is as follows.

> Accept Fact: so take any sentence S that is undecidable but understood by a speaker X. That is, suppose that X grasps the meaning of S, but possesses no means by which he can recognize either that S is true or that S is false. By Realism, either the condition for the truth of S obtains, or the condition for its falsity obtains. If the former, X nevertheless, *ex hypothesi*, cannot show that he recognizes the fact; if the latter, likewise. But now this contradicts Manifestation, which requires that X should be able to display his grasp of the meaning of the sentence X by the exercise of such a recognitional capacity concerning it.

At this point the opponent of the reductio might attempt to reply as follows:

> All that Manifestation requires is that grasp of the meaning of S be displayed in the exercise of *some* recognitional capacities concerning S; and this does not single out as necessary the capacity to recognize S as true if it is true, and to recognize S as false if it is false. Rather, the recognitional capacities could be exercised with regard to constituents of S, in contexts other than S itself, so that grasp of meaning of the various constituents

> of S could be severally displayed; whence, by compositionality, the meaning of S itself could be displayed as grasped.

This reply will not help, for one can counter it as follows:

> We shall be able to identify at least one constituent that is responsible for the undecidability of S. Let us suppose it is universal quantification over numbers. The universal quantification even of a decidable numerical predicate can be *pro tempore* undecidable; the classical meaning of the universal quantifier accordingly contributes a possibly recognition-transcendent ingredient to the meaning of any arithmetical sentence in which it occurs. Now such recognitional capacities as *can* be exercised in evidence of semantic understanding are, in the arithmetical case, always concerned with the status of texts as *proofs*. But, given the well-known incompleteness phenomena (which, it is worth noting, can be established intuitionistically) the semantic understanding whose manifestation is thus limited cannot support attributions to arithmetical statements of such meanings as can be relied upon in general to render them determinate in truth value.

Now McGinn claims in his Reply to my paper that I had misunderstood the role played, in his exposition of Dummett's argument, by his (McGinn's) claim (p. 21) that

> understanding a sentence ... consists in an ability to determine its truth value in some canonical way.

He assures us that what he was doing in claiming that was setting out the realist position before going on to expound the Dummettian reductio. All this I well appreciate; yet still my original criticism of McGinn holds.

Why? Because McGinn has not discerned sufficient structure in Dummett's argument to appreciate exactly what role the idea of canonical determination of truth value plays in his (McGinn's) own thinking. I intend to show just how he is still at root committed to the claim quoted above (even if only in a slightly generalized form) and therefore still subject to my original criticisms.

His quoted claim ill-advisedly conflates two theses that ought to be kept apart in a proper exposition of Dummett's reductio. The first of these theses is Realism itself, the proper focus of the reductio. The second is McGinn's mistaken version of Manifestation, which I shall call 'McGinn's version of manifestation':

> Understanding a sentence consists in an ability to determine in some canonical way whether or not it possesses the central property of the molecular semantic theory.

McGinn's exposition of the Dummettian reductio can now be seen to have McGinn's version of manifestation and Realism as separate hidden premisses for his quoted claim above. Overall the reductio, on this diagnosis of McGinn's presentation, is as follows:

> *McGinn's version of manifestation*: understanding a sentence consists in an ability to determine in some canonical way whether or not it possesses the central property of the molecular semantic theory
>
> plus *Realism*: the central property of sentences for molecular semantic theory is (classical) truth
>
> yields the *quoted claim*: understanding a sentence consists in an ability to determine its truth value in some canonical way;
>
> but *Fact*: we understand some *pro tempore* undecidable sentences, i.e. sentences whose truth value we cannot at present effectively determine –
>
> Contradiction!

The proper focus of Dummett's reductio is, as I have already said, Realism. If we are to focus properly on Realism as an assumption for the sake of such reductio then it seems to me not at all uncharitable to pin on McGinn that version of Manifestation that I have here called McGinn's version of manifestation. But then McGinn remains committed to the latter when arguing on behalf of Dummett by way of exposition. But it is precisely McGinn's version of manifestation that I was at pains to disavow as a supposed plank of the Anti-realist platform.

Why disavow it? For reasons both positive and negative. Firstly, one can give more sensible detail (as I tried to do in my paper) to Manifestation in terms of capacities to recognize purported proofs and disproofs as conclusive or not, rather than in terms of capacities to *find* proofs or disproofs (hence the title of my earlier paper). Secondly, if we assume McGinn's version of manifestation then the whole argument form above is easily adapted to provide a parallel reductio of the Anti-realist's claim. I am assuming that Dummett himself would have seen this, and that for that very reason it must be a mistake to visit McGinn's version of manifestation upon him as an undischarged assumption of an argument designed to reduce Realism to absurdity. Here is the adaptation in question:

> *McGinn's version of manifestation*: understanding a sentence consists in an ability to determine in some canonical way whether or not it possesses the central property of the molecular theory
>
> plus *Anti-realism*: the central property of sentences for molecular semantic theory is warranted assertability
>
> yields (*quoted claim*): understanding a sentence consists in an ability to determine in some canonical way a warrant for its

assertion or a warrant for its denial (i.e. for the assertion of its negation);

but (*Fact*): we understand some *pro tempore* undecidable sentences, i.e. sentences for whose assertion or denial we cannot (at the time we claim to understand them) effectively determine any warrant.

Contradiction!

McGinn's version of manifestation invites the old positivists' mistake: that of a three-fold division of sentences into the determinably true, the determinably false and the meaningless. It would be a mistake even if 'true' and 'false' were given anti-realist readings, as 'warrantedly assertable' and 'warrantedly deniable' respectively. Given bivalence, the classicist revises the trichotomy to the determinately true, the determinately false and the meaningless.

The anti-realist will accept neither of these tripartite schemes. His ascriptions of meaning are based on a molecular semantics. Such a semantics describes conditions of correct usage of basic expressions (both logical and extra-logical), and accounts for how meanings of complex expressions depend, in a way governed by their logical structure, on the meanings of their parts. In the description of basic meanings, extra-logical vocabulary will receive criterial treatment, with details depending on the area of discourse in question; while logical vocabulary will receive inferential treatment, with introduction rules playing a prominent role in a definition of validity of arguments. In the structural account of meaning-dependence, the anti-realist need do no more than borrow from Tarski and Davidson: for a disquotational theory of truth satisfying the adequacy condition (Convention T) is a disquotational theory of warranted assertability. (For a proof that this is so, see my paper 'Holism, Molecularity and Truth'.)

If one understands all the basic expressions in a sentence, and understands the way they are put together to form the sentence, then one understands the sentence. On this fundamental truth the molecular semantic method is based. The theory it leads to heeds the Manifestation requirement. For one can say what counts as displaying grasp of the meanings of basic expressions, and what counts as grasp of the way these expressions are put together in any sentence. In the case of names and atomic predicates, various kinds of observable behaviour will criterially warrant ascription of grasp of meaning to the language user. Examples are recognition and matching tests for colour predicates, and ability to perform basic computations in the case of arithmetical predicates (or functions). One can likewise say what sort of observable behaviour manifests grasp of the meaning of a logical operator. Take the example of negation. The language user can safely be said to grasp the meaning of negation if he both recognizes a reductio proof as warranting the denial of its assumption, and tries to warrant the denial of a sentence by taking it as the assumption in a reductio argument. Similar accounts can be given for the other logical operators. A speaker's proof-seeking and proof-checking behaviour can also warrant the claim that he understands the logical structure of the sentences involved. For the order in which he breaks assumptions and conclusion into immediate constituents when posing deductive sub-problems for solution on the way to solution of the main problem (be the problem one of discovery or of appraisal) will show whether he understands scope distinctions etc. – in short, whether he understands the way the sentences are composed out of their constituent expressions.

Proof searches, of course, may be unsuccessful; may, indeed, be so for a sentence and for its negation. But that is not to say that there is more to grasp of meaning than can be made manifest in the activities of searching for proofs and of checking purported proofs for proofhood. Indeed, I would still go so far as claiming that grasp of the meanings of logical operators can be fully manifested in just the latter activity, namely in proof checking alone. But even if one includes proof search as necessary further manifestation of grasp of meaning, molecularity ensures meanings for well formed wholes once given meaningful parts. Consider, for example, the case of a mathematician who for his first time is posed the problem of finding a proof of Goldbach's conjecture. That is, he is given

> RTP For every even number n greater than 2 there are prime numbers p, q such that n = (p + q).

Suppose his response is the following:

> Let n be an arbitrary even number greater than 2. So for some k greater than 1, n = 2k. I need to find primes p, q such that n = (p + q). Clearly I need only look among the primes less than n. Let's check some examples to see if there's a pattern that we can use in an inductive argument:
> 4 = 2 + 2
> 6 = 3 + 3
> 8 = 5 + 3 ... so we can't have p = q = n/2; anyway, that would be silly, because lots of odds aren't prime, and n/2 may be even in the general case ...
> 10 = 3 + 7
> 12 = 5 + 7
> 14 = 7 + 7
> 16 = 3 + 13 ... now here's a case where the only two primes that will do the trick differ by more than n/2 ... but wait! we also have
> 16 = 5 + 11 ... so could it be the case that primes can always be chosen that differ from each other by less than n/2? ...
> 18 = 7 + 11
> 20 = 3 + 17
> = 7 + 13 ...

> Nothing general seems to be emerging from this. I need to show that the prime sum can be done for $2(k+1)$ if it can be done for $2k$, but the pattern keeps jumping about. I can't simply add 2 to p or add 2 to q, because I don't in general obtain a prime by doing so.
>
> I wonder, indeed, whether there isn't a counterexample? Say, who is Goldbach anyway? Could *he* prove it? What?? Do you mean to tell me that you've set me a problem that even famous mathematicians haven't been able to prove? That's not fair. I mean, I understand what you want me to prove, but you can't really expect me to prove it.

Our mathematician proves himself easily frustrated and mediocre, but his complaint is well taken. He has displayed full grasp of the meaning of Goldbach's conjecture, but has failed in his search for a proof, as everyone has before him. There may be more to mathematics than meaning, but there is no more to mathematical meaning than the basic code of mathematical practice.

Goldbach's conjecture is undecidable. By this I mean not that it is impossible to prove and impossible to refute; I mean, rather, that we do not at present possess (an effective method for finding) either a proof or a disproof of it. If 'undecidable' had the stronger reading I have rejected, one would never be able to warrant an assertion to the effect that a sentence was undecidable. For to do so would be in part to warrant the assertion that it was impossible to prove. This in turn would involve reducing to absurdity the assumption that one could prove it. But this would be to disprove it, hence decide it. It is clear that the weaker reading just given is the one involved in my claim that

> The anti-realist can admit the possibility of definitely meaningful but undecidable sentences.

Goldbach's conjecture, though undecidable, is meaningful. Our mathematician, moreover, grasped its meaning. Thus Weir's criticism (p. 178) that

> to refute the view that an undecidable sentence has no meaning does not imply that the sentence has a determinate meaning

simply misses the mark in this case. To be fair to Weir, however, it should be noted that his criticism was levelled at an account of manifestation of grasp of meaning via proof checking rather than proof search. To meet this criticism squarely, therefore, let us confine ourselves to proof checking, when assessing the recognitional abilities of a speaker that are, on the anti-realist's account, criterial for his grasp of meaning. Weir says (p. 178) that

> abilities of the sort Tennant alludes to – abilities to decide proofhood of purported proofs – certainly constitute ground for denying the denial of meaningfulness. But do they constitute grounds for asserting that the sentence has a determinate meaning?

Of course they do. I am at a loss to understand why Weir should, on behalf of the anti-realist, think otherwise. Is he raising a general scepticism about rule following? Is he saying that we cannot assert that a speaker has mastered the meaning of conjunction, say, if all we have is a finite amount of evidence concerning his apparently correct usage of the conjunction sign in past inferences? and is he supposing that this possible indeterminacy in the meaning attached by the speaker to the conjunction sign will infect any compound in which the sign occurs? If this is Weir's misgiving then it is one which has to be allayed by any behaviouristic theory of meaning, realist or anti-realist. I cannot undertake to allay that misgiving here. If, on the other hand, Weir is not a grasp-of-meaning sceptic for atomic expressions, he must be locating indeterminacy of sentence meaning (for the speaker) in an alleged indeterminacy (again, for the speaker) of logical structure in the sentence concerned. But here again the same sort of evidence is available to ground attributions of correct logical diagnosis. We can see how the speaker partially orders his subroutines in checking proofs (and searching for them); how he breaks assumptions and conclusion into their immediate constituents; how he is sensitive to scope distinctions by invoking rules of inference in the right partial order. And once given determinacy of atomic meanings and determinacy of logical structure, there is no further source of indeterminacy of sentence meaning. Weir himself adopts a molecular account along these lines when he outlines his own brand of behaviourist but realist semantics: for even undecidable sentences

> will be semantically complex, constructed out of simpler components by means of sentential, and other, operators. *If determinate extensions*, such as truth functions, *can be associated with such operators* then, given that the simplest components have determinate truth values, so too will the complex sentences [my emphasis].

But what about the language of arithmetic? The classicist and intuitionist agree on all quantifier free sentences of arithmetic. Such sentences are decidable, hence irrelevant to locating the source of disagreement between the classicist and the intuitionist. Their disagreement hinges on the interpretation of quantifiers. Weir would have to allow us to read further with the changes indicated in his text:

> Understanding of the (interpretations) of the (quantifiers) will, on a behaviouristic theory, be a matter of appropriately patterning responses to complex sentences with the (quantifier) dominant, given responses to the constituents.

What, then, is the appropriate pattern of responses to a quantified sentence, given what responses to what constituents? A positive response to (a purported proof of) F(t) warrants positive response to (because in one step we can extend the given proof to a proof of) $\exists x F(x)$. A negative response to F(t) likewise warrants a negative response to $\forall xF(x)$. (Here I take a negative response to $F(t)$ to be a positive response to a purported disproof of $F(t)$.) But it is crudely fallacious to regard positive responses to (purported proofs of) any $F(t1), \ldots, F(tn)$ as warranting positive response to the universal claim $\forall xF(x)$. One simply cannot be done surveying responses to $F(0)$, $F(1)$, $F(2), \ldots$ *ad infinitum*. One would have to look rather for positive response to a form of argument involving the constituent predicate F, schematically:

$$\dfrac{F(0) \quad \begin{array}{c}\overline{F(n)}\\ \vdots \\ F(n')\end{array}}{\forall xF(x)}$$

Weir's behaviourist account of how truth values are determined on the basis of values of constituents simply breaks down when applied to universal quantifications. His own behaviouristic method will oblige him to look elsewhere for the sort of behaviour that warrants ascription of grasp of meaning to the speaker. I am suggesting that the only place to look, in the nature of the case, is applications of the rule of mathematical induction. Same behaviouristic method, but (dare I say?) more realistic choice of behaviour. The behaviour is all that grounds our understanding of the language and its subject matter; and proper attention to the implications of such a publicity principle forces one to anti-realist conclusions about meaning.[1]

University of Stirling,
Stirling, FK9 4LA

© NEIL TENNANT 1984

REFERENCES

Colin McGinn, 'Truth and Use', in Mark Platts (ed.), *Reference, Truth and Reality* (Routledge, 1980), 19–40.

Neil Tennant, 'Is This a Proof I See Before Me?', *Analysis*, 41.3, June 1981, 115–19.

Colin McGinn, 'Reply to Tennant', ibid., 120–2.

Neil Tennant, 'Holism, Molecularity and Truth', in B. Taylor (ed.) *Festschrift for Michael Dummett* (Nijhoff, forthcoming).

Alan Weir, 'Truth Conditions and Truth Values', *Analysis*, 43.4, October 1983, 176–80.

---

[1] I am indebted for their comments on earlier versions of this paper to Crispin Wright and Barry Smith.

SAINSBURY ON A FREGEAN ARGUMENT

By STIG ALSTRUP RASMUSSEN

IN a recent paper (‘On a Fregean Argument for the Distinctness of Sense and Reference’, ANALYSIS 43.1, January 1983) R. M. Sainsbury argues that Frege’s celebrated argument in support of the need for distinguishing between sense and reference is inconclusive. Sainsbury recasts the Fregean argument as follows:

(1) a = b
(2) ‘a = a’ is *a priori*.
(3) ‘a = b’ is not *a priori*.
(4) If s(e) = s(e′), then ‘. . . e . . .’ is *a priori*, iff ‘. . . e′ . . .’ is *a priori*.
(5) For all x and y, if x$^\frown$‘=’$^\frown$y is true, then r(x) = r(y).

Hence, by *reductio*

(6) Not: for all x, s(x) = r(x).

The expressions ‘a’, ‘b’, ‘r(x)’ and ‘s(x)’ abbreviate ‘Hesperus’, ‘Phosphorus’, ‘the reference of x’ and ‘the sense of x’ respectively. The above argument does not, of course, regiment everything Frege has to say, in the opening pages of ‘Über Sinn und Bedeutung’, in support of his distinction, since, for instance, it takes no account of his rebuttal of the metalinguistic construal of identity sentences. Also, as Sainsbury points out, (6) is considerably weaker than Frege’s conclusion, which is that the sense and the reference of an expression, as it occurs within any particular context, never coincide. However, these omissions tend, if anything, to make Sainsbury’s criticism all the more alarming to the Fregean.

Sainsbury deems the Fregean argument inconclusive on the ground that there are ways of disputing (3) and (4). First, on the supposition that ‘a = b’ and ‘a = a’ express the same proposition, the adoption of

(I) σ is *a priori* iff anyone who understands σ can, solely in virtue of this understanding, come to know the truth of the proposition which σ expresses (p. 13),

constrains us, Sainsbury says, to give up (3). Secondly, if we endorse

(II) σ is *a priori* iff anyone who understands σ can, solely in virtue of this understanding, come to know that σ expresses a true proposition (p. 13),

Sainsbury considers us committed to abandoning (4). He concludes that in order to render the Fregean argument persuasive ‘we would have to dig deeper into such notions as *a priori*, *proposition* and *understanding*’ (p. 14).

I shall argue that Sainsbury's conclusion is unduly pessimistic: while it is certainly true that to get a firm grasp of the content of the Fregean notion of sense, we would have to undertake a deeper enquiry of the sort he adverts to, nevertheless the Fregean argument in itself amounts to an overwhelmingly powerful case for the need to introduce a distinction between sense and reference, provided (4) is divested of some of its generality. In addition, even if the Fregean had to concede Sainsbury's criticism, he could easily reshape his argument in a manner such as to make it immune to the objection.

My reservation concerning (4) is this. It is obvious that, prior to specifying just what the sense of an expression is, the Fregean is in no position to help himself to a sweeping principle like (4). Fortunately, however, his argument does not require anything stronger than

(4′) If s(e) = s(e′), then 'a = e' is *a priori* iff 'a = e′' is *a priori*, where the expressions flanking the identity signs are ordinary proper names.

The substitution of (4′) for (4) leaves unaffected Sainsbury's objection; but in view of the fact that (4) is genuinely controversial for reasons having nothing to do with the objection, it is preferable to consider the argument in the form in which it involves (4′).

Now, Sainsbury's argument to the effect that there is a way of denying (3) appears to run as follows:

(7) Concerning the proposition expressed by 'a = a', we may, solely in virtue of understanding 'a = a', come to know that that proposition is true (from (I) and (2)).
(8) The proposition expressed by 'a = a' is identical to the proposition expressed by 'a = b'.
(9) Concerning the proposition expressed by 'a = b', we may, solely in virtue of understanding 'a = a', come to know that that proposition is true (from (7) and (8)).
(10) Understanding 'a = b' entails, or involves, understanding 'a = a'.
(11) 'a = b' is *a priori* (from (9), (10), and (I)).

It deserves note that (10) is absolutely essential to Sainsbury's argument, since (9) does not suffice to prove that 'a = b' is *a priori* in the sense of (I). Since (10) looks, perhaps, a bit fishy, it deserves to be pointed out, too, that Sainsbury might defend it by saying that in order to understand 'a = b' one clearly has to understand what is involved in equating objects; and, arguably, understanding sentences like 'a = a' is part of what is involved in this.

When confronted with this argument, one immediately wants to protest that, surely, 'a = b' cannot be *a priori*. There is no way we could have established that 'Hesperus = Phosphorus' is true save by undertaking empirical enquiries; and that is what we intend to say when denying that the sentence is *a priori*. Sainsbury of course does not suggest that we might have established the truth of that sentence independently of empirical investigations. On the contrary, he offers an explanation why 'a = b' *appears* not to be *a priori*:

> What gives rise to apparent absence of *a prioricity* is the possibility of not knowing, concerning a sentence one understands and a proposition which one knows, that the one expresses the other (p. 13).

That is, understanding a sentence does not entail knowing what proposition it expresses. The suggestion is, then, that the empirical investigations required for establishing the truth of 'a = b' might be interpreted as establishing that 'a = b' expresses the same proposition as does 'a = a'. One would still like to know, however, why we should wish to say that a sentence is *a priori* even if we cannot establish its truth without engaging in empirical research.

An additional worry is this. It is not at all clear why, on the view under consideration, the otherwise trivial (7) should be acceptable. When explaining how the appearance that 'a = b' is *a posteriori* can arise, Sainsbury has the interlocutor saying that understanding a sentence does not, in general, suffice for knowing which proposition it expresses. However, there is little doubt that the only reason why Sainsbury thinks the Fregean's interlocutor entitled to (7) is that, in this special case, understanding the sentence involves knowing which proposition it expresses, combined, of course, with the fact that it is trivially true that a = a. So, we need to be told the difference between the cases where understanding a sentence suffices for knowing which proposition it expresses, and the cases where this does not hold. The suspicion is that the former cases will turn out to be those in which the sentence is *a priori* in sense (II); while the latter cases will be those in which the sentence is not. However, if we need *a prioricity* in a sense akin to the one specified in (II) anyway, it is difficult to see what the merits of (I) could be.

The Fregean, then, may be excused for thinking that the possibility of contesting (3) does not seriously affect the persuasiveness of his argument. How do matters stand in respect of (4′)? Sainsbury concedes that if we adopt (II) as our account of the *a priori*, (3) is true. The worry now is, he says, that (2) and (3) combine to invalidate (4′):

> Putting 'b' for one occurrence of 'a' in 'a = a' transforms the sentence into one such that one cannot know, solely on the basis of understanding it, that it expresses a truth (p. 13).

This is true, of course; but (2) and (3) do not suffice to prove (4′) false. We need, in addition, the premiss that the sense of 'a' equals that of 'b'; and Sainsbury fails to say why this should be supposed to be true. Two possibilities come to mind. Sainsbury may think himself entitled to assume s('a') = s('b') on the strength of the fact that the reference of 'a' is identical to that of 'b'; and the presupposition that sense equals reference. If so, however, he clearly begs the question, since the point at issue is precisely whether or not the presupposition holds. Alternatively, the identification of the proposition expressed by 'a = a' and the proposition expressed by 'a = b' is still in play; and Sainsbury takes s('a') = s('b') to follow from that identification. In this case, the Fregean might well grant the validity of the inference; but he would be right to complain that the identification of the propositions expressed by 'a = a' and 'a = b' is little short of question-begging. No one would accept s('a') = s('b') who did not accept that the sense of a proper name may be identified with its reference; so, in the presence of the inverse principle of functionality, the identification of the proposition expressed by 'a = a' with the one expressed by 'a = b' would rightly be rejected offhand by the Fregean. It deserves note that if Sainsbury takes the interlocutor to endorse the principle in question, the point just made holds, too, for his discussion of how to prove (3) false on the basis of (I).

The upshot so far is this. Sainsbury's argument from (I) and the identification of the proposition expressed by 'a = a' and the one expressed by 'a = b' to the repudiation of (3) threatens to render pointless the distinction between the *a priori* and the *a posteriori*. At the same time, the description of the position from which he argues seems capable of completion only if another notion of the *a priori* is invoked, viz. some notion closely akin to the one defined in (II). Finally, the argument may be little short of question-begging. As regards Sainsbury's argument from (II) to the falsity of (4′), it is incomplete as it stands; and when completed, the argument is either explicitly question-begging or falls little short of being so.

However, the Fregean does not have to rely on (1)–(6). It is true, of course, that when presenting his argument in support of the need for distinguishing between the sense and the reference of a proper name, Frege invokes the notion of the *a priori*; but he indicates that the notion of cognitive value ('Erkenntniswert') would do just as well. ('On Sense and Meaning', in P. Geach and M. Black (eds), *Translations from the Philosophical Writings of Gottlob Frege*, Oxford 1980, p. 56.) Suppose, then, (2) and (3) are replaced by

(13) 'a = a' and 'a = b' differ in respect of cognitive value;

and we substitute for (4′)

(14) If s(e) = s(e′), then the cognitive value of 'a = e' is identical with that of 'a = e′'.

It seems impossible to deny (13), and (14) may justly be construed as a constraint on the notion of sense. The Fregean might argue as follows. Suppose we agree that (13) holds. Then, surely, we need some notion in terms of which the difference in respect of cognitive value of 'a = a' and 'a = b' can be explained. If this is the job we have in mind for the notion of sense, (13) ensures that the sense of 'a' differs from that of 'b'. If so, however, (14) must be a valid principle. It follows that sense and reference must be distinguished, since 'a' and 'b' refer to the same object.

Now, Sainsbury does not directly discuss this version of the Fregean argument; but he indicates that what might be taken to be a precisification of it fails (the Editor pointed out to me that Sainsbury's (3″) and (4″) might be viewed thus):

> Or one might prefer, in place of (4),
>
> (4″) if s(e) = s(e′), then 'A believes that ... e ... ' is true iff 'A believes that ... e′ ... ' is true.
>
> This would return us to a situation parallel to the one produced by (4). The defender of [the negation of (6)] may accept that
>
> (3″) 'A believes that a = a' is true iff 'A believes that a = b' is true
>
> and use this to show the falsehood of (4″); or he may accept (4″) and use this to deny (3″) (p. 14).

Two preliminary points. In the first place, the Fregean does not need anything so sweeping as (4″). He can rest content with

(14′) If s(e) = s(e′), then 'A believes that a = e' is true iff 'A believes that a = e′' is true.

Secondly, Sainsbury obviously has in mind not (3″), but the negation of that statement (I owe this point to the Editor):

(13′) 'A believes that a = a' and 'A believes that a = b' may take different truth values.

Now, how could (13′) possibly fail? According to Sainsbury, there ought to be an argument from (14′) to the falsity of (13′); but it seems that there could be no such argument, unless s('a') = s('b') can be taken for granted. This, however, it cannot, for reasons pointed out above. In fact, (13′) seems absolutely unassailable. As regards (14′), the Fregean is entitled to reason in a manner analogous to the reasoning advanced in support of (14): since (13′) is evidently true, there must be some difference between 'a = a' and 'a = b' which accounts for the fact that one might simultaneously believe the one and disbelieve the other. Since the only difference between the two sentences is that 'b' occurs in one of them at a place where 'a' occurs in the other, the difference must be explained in terms of some property of ordinary proper names. Suppose, then, we say that 'a' and 'b' differ in respect of sense. If this is how the notion of sense is introduced, (14′) cannot fail.

I am not suggesting that the above is an adequate description of Frege's reasons for introducing the notion of sense. Evidently, Frege considered the notion of sense subject to additional constraints. However, I have refrained from appealing to the connection between the sense of an expression and what understanding the expression consists in, because there seems to be no way of doing so that does not run the risk of begging the question against the Fregean's interlocutor, as this character is depicted by Sainsbury. The fact remains that appealing to (13′) and (14′), and basing the latter on considerations such as the above, is *a* way of making good the claim that there is a need for distinguishing between the sense and the reference of a proper name.[1]

University of Copenhagen/
Wolfson College, Oxford
OX2 6UD

© STIG ALSTRUP RASMUSSEN 1984

---

[1] I am indebted to Michael Dummett, Jens Ravnkilde, and the Editor for comments and suggestions.

REJOINDER TO RASMUSSEN

By R. M. SAINSBURY

THE strategy in the paper of mine which Stig Alstrup Rasmussen discusses[1] was to consider how persuasive is a certain Fregean argument by seeing what one would be committed to if one were to deny its conclusion. I considered whether one who held that there is no distinction between sense and reference for names, or, better, that names with the same reference have the same sense, would be committed, by the Fregean argument, to obvious absurdities.

Among other things, Rasmussen accuses me of begging a question against Frege; and he claims that a variant of the argument I discuss would constitute an 'overwhelmingly powerful case' for the Fregean conclusion, while being immune to my objections.

The original version of the argument I consider turns upon the concept of *a priority*. I gestured towards two possible accounts of this concept, and showed, with respect to each, that one who held that sameness of reference is enough for sameness of sense, and so, in particular, that s('a') = s('b'), is not shown by the Fregean argument to be committed to obvious absurdity. (The accounts of *a priority* are not endorsed: they are intended partly as a challenge, and partly to illustrate a general difficulty which will beset any concept designed to play that role in the Fregean argument.) Rasmussen's charge of circularity relates to my attempt to show this for the second account of *a priority*:

---

[1] My original paper, 'On A Fregean Argument for the Distinctness of Sense and Reference' appeared in *Analysis* 43.1, January 1983. Rasmussen's discussion, 'Sainsbury on a Fregean Argument', appears in this issue.

> σ is *a priori* iff anyone who understands σ can, solely in virtue of this understanding, come to know that σ expresses a true proposition.

To rebut the charge of circularity, let me state more fully the reply I envisage on behalf of one who holds that s('a') = s('b'). He will accept that

(2) 'a = a' is *a priori*

and that

(3) 'a = b' is not *a priori*.

But he will reject

(4) if s(e) = s(e′), '. . . e . . .' is *a priori* iff '. . . e′ . . .' is *a priori*.

Indeed, he will see (2) and (3) as grounds for his rejection of (4), and, bearing in mind the account of *a priority*, this rejection is highly plausible. As applied to 'a' and 'b', (4) has a true antecedent, on the present assumption; but (2) and (3) entail the falsity of its biconditional consequent. The assumption of s('a') = s('b') is not a begging of the question against Frege, but an implementation of the strategy of seeing whether the denial of the Fregean conclusion is shown by the argument in question to lead to obvious absurdities.

The most that this shows is that the Fregean argument is not well served by the concept of *a priority* as specified in my second account. So greater interest attaches to Rasmussen's claim that replacing this by another concept provides a persuasive argument. He suggests, in effect, that my (2), (3) and (4) be replaced by, respectively:

(2R) X believes that a = a
(3R) X does not believe that a = b
(4R) If s(e) = s(e′), then ('X believes that a = e' is true) iff ('X believes that a = e′' is true).

In the original paper, I suggested (perhaps too briefly[2]) that one who wants to hold that s('a') = s('b') could respond in either of two ways. He could, without obvious absurdity, deny (3R). Sincere denials of the sentence 'a = b' would be interpreted not as showing that the utterer does not believe that a = b, for the anti-Fregean will naturally see this as equivalent to not believing the trivial proposition that a = a. Rather, the denials will show merely that the utterer had not grasped what proposition the sentence 'a = b' expresses. And the apparent incorrectness of saying, of such a person, that he believes that a = b will be regarded as misleadingness, to be explained by conversational factors: such ascriptions of belief will be strictly true. (The general idea, that what is believed (= proposition believed) can be identified by its truth-conditions, and divorced from what one can identify merely on the strength of understanding a sentence, is taken seriously (I do not say endorsed) by, for example, David Kaplan in 'Demonstratives', John Locke Lectures, Oxford University, 1980.)

---

[2] In addition to excessive brevity, my original discussion contained a mistake: its (3″) should have been negated. I apologize; and I thank Rasmussen and the Editor for the correction.

Alternatively, the adherent of the view that s('a') = s('b') might cling to both (2R) and (3R) but deny (4R). Its falsity would be made manifest by the truth of (2R) and (3R). (In the original paper, I pointed out that David Wiggins has given us independent reason to suppose that sameness of sense is not a sufficient condition for substitution *salva veritate* in all contexts.) My claim in the original paper was that neither option is patently absurd. To reject either, we would have to dig deeper into the notions of *proposition*, *understanding*, and so on.

I would now see a general moral of my paper as being close to a point emphasized by Daniel Dennett (in 'Beyond Belief', in Andrew Woodfield, ed., *Thought and Object*, pp. 1-95, at pp. 6-10 and ff.). Frege wants sense to determine thoughts, and thoughts, for Frege, have two properties: they are what is grasped in linguistic mastery, and they are truth bearers. The Fregean argument I discussed should be persuasive only to someone who was already persuaded that these properties must be possessed by a *single* entity. Some of the moves I offer to the adherent of the view that s('a') = s('b') consist in forcing these properties apart, in opening a gap between what is yielded by linguistic mastery, and what is the bearer of truth value; or, in the terminology of the original paper, between understanding a sentence and knowing what proposition it expresses. The paper was neutral on the question of whether the unification effected in Frege's conception of a thought is correct or incorrect. But this unification needs to be in play at an earlier stage in 'On Sense and Reference' than is usually recognized: for only if it is assumed is the argument I purport to extract from the opening paragraphs of that paper persuasive.

Bedford College
(University of London),
Regent's Park, London NW1 4NS

© R. M. SAINSBURY 1984

IMPOSSIBLE WORLDS

By NATHAN SALMON

IN a recent commentary on my *Reference and Essence* (Princeton University Press and Blackwell, 1982), William R. Carter represents me as endorsing the first two of the following three modal propositions, which together constitute an inconsistent triad (following Carter's numbering scheme):[1]

(2) If a given table (ship, bicycle, etc.) x is originally constructed from a certain hunk of matter y, then it is a necessary or essential feature of x that it is originally constructed from y.

(3) If a given table (ship, bicycle, etc.) x is originally constructed from a hunk of matter (collection of material parts) y, then x could have originated from a hunk of matter z 98% (or more) of which overlaps with y; but x could not have originated from any hunk of matter z', such that less than 98% of z' overlaps with y.

(4′) If c is a material component (e.g., a molecule) of a hunk of matter y, then it is a necessary or essential feature of y that it has c as a material component.

In fact, I endorse (4′), but neither (2) nor (3). The strongest principle along the lines of (2) that I endorse is the following:[2]

(2′) If a given table (ship, bicycle, etc.) x is originally constructed from a certain hunk of matter y, then x could have been originally constructed from any hunk of matter z which is sufficiently like y (in mass, volume, composition, etc.) and which sufficiently substantially overlaps y; but x could not have been originally constructed from any hunk of matter z' which does not sufficiently substantially overlap y.

I offer (3) as one among uncountably many possible regimentations or sharpenings of (2′), one candidate for what is to count as *sufficiently substantial* overlap. I do not actually endorse (3), however, since I regard the vagueness of the phrase 'sufficiently substantial overlap' in (2′) as intrinsic to the epistemic situation.

---

[1] Salmon on Artifact Origins and Lost Possibilities', *The Philosophical Review*, XCII, No. 2 (April 1983), pp. 223–31.

[2] Here I assume the following modal evaluation clause:

> $\ulcorner \Box\phi \urcorner$ is true with respect to w iff ϕ is true with respect to every world determinately accessible to w and either true or neither true nor false with respect to any world neither determinately accessible nor determinately inaccessible to w.

For more on indeterminate accessibility, see *Reference and Essence*, pp. 247–52. The evaluation clause assumed here differs from the (strong) rule proposed there at p. 248, note 27.

No precise principle, like (3), which removes the vagueness by substituting sharp cut-off points is knowably true. (Cf. *Reference and Essence*, pp. 240-52.)

A principle like (2′) is not the sort of proposition that merely happens to be true. If it is true at all, it is necessarily so. In fact, if it is true at all, then it is necessary that it is necessarily true, and it is necessary that it is necessary that it is necessarily true, and so on. From this observation, a sorities-type construction, the main idea of which has been exploited by Roderick Chisholm,[3] can be made to show that the generally accepted axiom schema of *S4* modal propositional logic,

$$\Box p \supset \Box\Box p$$

or equivalently, the notion that modal accessibility among worlds is transitive, should be rejected in its unrestricted form. Consider any hunk of matter z which is sufficiently like the hunk of matter y (in mass, volume, composition, etc.) that originally constitutes an artefact x, but which does not sufficiently overlap y. By (2′) it is necessary that artefact x is not originally formed from hunk z. But there is a (perhaps scattered) hunk of matter z_1 which includes some of the molecules of hunk z and which does sufficiently overlap hunk y, so that artefact x could have been formed from hunk z_1. Consider now yet another hunk of matter z_2 which includes still more of the molecules of hunk z and which sufficiently overlaps hunk z_1 (though perhaps does not sufficiently overlap hunk y). If artefact x could have been formed from hunk z_1, then (even if, in fact, x could not have been formed from z_2, still) it *might have been* that x could have been formed from z_2. Continuing in this vein, it will follow that, although it is necessary that artefact x is not formed from hunk z, still it might have been that it might have been that it might have been ... that x is formed from z. More intuitively, if there is a possible world w_1 (possible relative to the actual world) in which artefact x is formed from hunk z_1, then there is a world w_2 possible relative to w_1 in which x is formed from z_2. Hence there is a world w_3 possible relative to w_2 in which artefact x is formed from a hunk of matter z_3 which includes still more molecules of hunk z and which sufficiently overlaps hunk z_2, and so on. Finally, there will be a world w which bears the ancestral of the accessibility relation to the actual world, and in which artefact x is formed from hunk z, though by hypothesis there is no world accessible to the actual world in which x is formed from z. World w is an *impossible world* from the point of view of the actual world.

---

[3] See his 'Identity Through Possible Worlds: Some Questions', *Noûs* 1 (March 1967), pp. 1-8; and *Person and Object* (La Salle: Open Court, 1976), pp. 148-9. Chisholm does not accept my conclusions concerning his argument, but instead rejects any principles like (2′) and (3) in favour of the inflexible essentialism of (2). See also Saul Kripke, *Naming and Necessity* (Harvard University Press and Basil Blackwell, 1980), at p. 51, note 18; and Hugh S. Chandler, 'Plantinga and the Contingently Possible', *Analysis* 36.1, January 1976, pp. 106-9.

Though the artefact x could not have been formed from hunk z, there is no reason why hunk z could not have been formed instead of hunk y into an artefact of exactly the same type and form as x in place of x itself. Thus there is a world w' possible relative to the actual world in which an artefact x', qualitatively just like x, is formed from hunk z. World w' is, we may suppose, materially exactly like the impossible world w in every molecular, atomic, and sub-atomic detail. Given a complete accounting of the entire history of all of the matter in the worlds w and w', with its causal interconnections and exact configuration through time, the two worlds are absolutely indistinguishable. Atom for atom, quark for quark, they are exactly the same. Yet they must be distinct, since w' is, and w is not, a *genuinely* possible world, i.e., a world possible relative to the actual world. (Cf. *Reference and Essence*, pp. 230–40.)

Carter objects to this 'model of the situation' on the grounds of a principle of the identity of materially indiscernible worlds. If the phrase 'materially indiscernible' is understood in such a way that w and w' count as materially indiscernible, then what we have here is an example which gives the lie to this principle. It is important to notice that w and w' are indeed discernible, even if not materially discernible in this sense, and in fact discernible not only by their accessibility relations to the actual world. They also differ as regards which facts obtain in them. World w includes the fact that artefact x is formed from hunk z, whereas world w' excludes this. Some other artefact x', distinct from x, is formed from hunk z in w'. In place of Carter's principle, I propose a principle of the identity of factually indiscernible worlds, worlds in which the very same facts obtain. (Cf. *Reference and Essence*, p. 238.) I also propose a principle of the identity of *mutually accessible* materially indiscernible worlds. (Cf. *Reference and Essence*, p. 240, and p. 249, note 28.) But an unbridled principle of the identity of materially indiscernible worlds is refuted by the sorts of considerations raised here.

Is this picture of impossible worlds and mutually inaccessible materially indiscernible worlds really acceptable? There are a number of conceptions of possible worlds presently in vogue. Possible worlds are variously construed as maximal compossible sets of propositions (Robert Adams), possible total histories or states of the world (Saul Kripke, Robert Stalnaker), total ways things or the world (cosmos) might have been (David Lewis, sometimes), maximal states of affairs (Alvin Plantinga). For present purposes, these need not be regarded as competing conceptions of possible worlds (except in the case of Lewis, who usually takes nonactual possible worlds to be something like immense concrete objects, someplace far, far away). On any of these conceptions, whatever grounds there may be for believing that there really are possible worlds yield the same, or related, reasons for believing that there are impossible worlds (maximal consistent though not compossible sets of propositions, impossible total histories of the world, impossible total states of the cosmos, total ways things could not have been, etc.), for believing that there are materially indiscernible worlds (materially indiscernible total histories of the cosmos, materially indiscernible total ways for things to be, etc.), for believing the identity of factually indiscernible worlds, and the rest.[4]

University of California,
Riverside, California 92521, U.S.A.

© NATHAN SALMON 1984

---

[4] I am grateful to the Editor of *Analysis* for pointing out a significant oversight in the original draft of this note.

INDEFINITE IDENTITY: A REPLY TO BROOME

By H. W. NOONAN

IN 'Can there be vague objects' (ANALYSIS 38.4, October 1978) Gareth Evans argued that there could be no objects about which it was a *fact* that they had fuzzy boundaries, and consequently no identity statements which were indeterminate in truth-value because one or both of their terms denoted such an object. In my 'Vague objects' (ANALYSIS 42.1, January 1982) I endorsed Evans' argument but pointed out that it left open the possibility of identity statements which were indeterminate in truth-value because one or both of the terms flanking the sign of identity lacked a determinate denotation. I then went on to consider some consequences of this being the *only* way in which identity statements could be indeterminate in truth-value, consequences which seemed to me to be interesting particularly in relation to the idea that there might be indeterminate statements of diachronic personal identity.

In 'Indefiniteness of Identity' (ANALYSIS 44.1, January 1984) John Broome challenges Evans' argument. In what follows I defend Evans' argument against Broome's criticisms and argue that the example he produces in support of his own conception of indefinite identity is in fact better understood in accordance with the account of indeterminate identity put forward in my previous paper.

A crucial point in Evans' argument (reproduced by Broome) is an application of Leibniz's Law. He argues that if '$a = b$' is indeterminate in truth-value (where this is a consequence of the fuzziness of *a* or *b* and not vagueness in either of the terms '*a*' and '*b*') there will be a property '$\hat{x}[\nabla(x = a)]$' (where '∇' is a sentential operator expressing indeterminacy) *b* possesses (i.e. definitely possesses) but *a* does not. Consequently, he argues, applying Leibniz's Law, the assumption that '$a = b$' is indeterminate in truth-value in this way leads to the conclusion that it is false, and so not indeterminate in truth-value. Hence '$a = b$' cannot be thus indeterminate in truth-value.

Broome objects to this argument that one cannot deduce that '$a = b$' is *false* from the fact that the predicate '$\nabla (x = a)$' is determinately true of *b* while its negation is determinately true of *a*, and that it is begging the question to assume that Leibniz's Law licences such a deduction.

The proper response for a defender of Evans' argument to this objection of Broome's, I think, is not to dispute with him about the proper formulation of Leibniz's Law, but to draw in his horns a little. Broome, he should admit, has shown that Evans' argument does not demonstrate the incoherence of *every* conception of indefinite identity. But it remains that Evans *has* demonstrated the incoherence of *a* notion of indefinite identity, and that the most natural one.

Asked to formulate an account of indefinite identity, I think one's first thought would be something like this. '$a = b$' is definitely true just in case any predicate definitely true of *a* is definitely true of *b*, while '$a = b$' is definitely false just in case there is some predicate definitely true of *a* whose negation is definitely true of *b*. If '$a = b$' is indeterminate in truth-value, then, there will be no predicate definitely true of *a* whose negation is definitely true of *b*, but it will not be the case that every predicate definitely true of *a* will also be definitely true of *b*; rather there will be predicates definitely true of *a* which are neither true nor false of *b*. What Evans' argument demonstrates is that this natural conception of indefinite identity can have no application: in the case in which '$a = b$' is indeterminate in truth-value (if this is not due to vagueness of '*a*' or '*b*'), just as in the case in which it is definitely false, there will be a predicate ('$\nabla(x = b)$') definitely true of *a* whose negation is definitely true of *b*.

The conception of indefinite identity which Broome has shown that Evans' argument does not refute, then, is in fact a conception of indefinite identity as a kind of *relative* identity: a relation which ensures indiscernibility of its terms in some, but not all, respects – in particular, not in respects only expressible by predicates containing '∇' or synonyms of such predicates. But I do not think that Evans envisaged his target in this way.

Of course this is unimportant if, as Broome claims, we can give a perfectly unpuzzling and transparent description of cases in which we are tempted to claim that certain identity statements are indeterminate in truth-value by using his notion of indefinite identity. But

I do not think that we can do this. On the contrary, applying Broome's notion to his own example of the club leads us, I suggest, to conclusions which it is hard even to understand, much less believe.

This example, described in as neutral a way as possible, is the following. Some people get together and engage in clublike activities for a certain period, say five years, at the end of which a certain event, which may or may not signal the demise of the club, occurs. For a certain period thereafter, say a year, no clublike activities take place, but then some of the original group, perhaps with some others, get together again and once more engage in clublike activities, and this continues, let us say, for a further twenty-five years. Now, as Broome says, it is tempting to describe a case of this sort, suitably elaborated, as one in which the facts render the identity statement 'the club existing earlier is the club existing later' indeterminate in truth-value. The question is whether a transparent and unpuzzling description of the case can be got by regarding the source of this indeterminateness as the holding between the objects determinately denoted by the terms 'the earlier club' and 'the later club' of Broom's relation of indefinite identity. I think not.

For if 'the earlier club is the later club' is indeterminate in truth-value for this reason, then the predicate 'lasted for at most five years' will be neither determinately true nor determinately false of the object determinately denoted by the term 'the earlier club' (for if the identity *were* false that predicate would be true of the earlier club, and if the identity *were* true it would be false of the earlier club). Similarly, the predicate 'lasted for at least twenty five years' will be neither determinately true nor determinately false of the earlier club (for if the identity were true that predicate would be true of it, and if the identity were false it would be false of it). On the other hand, the predicate 'lasted for at most five years or lasted for at least twenty five years' must be determinately true of the earlier club (for the object determinately denoted by 'the earlier club' has certainly lasted for at least five years on Broome's account, and there is no other longer-lived entity, apart from the later club – which has lasted for at least twenty five years – with which it might be identical, so its lifespan must either be a maximum of five, or a minimum of twenty five, years). The earlier club, then, on Broome's account, must be an object which determinately satisfies the predicate 'lasted for at most five years or lasted for at least twenty five years' but neither determinately satisfies the predicate 'lasted for at most five years' nor determinately satisfies the predicate 'lasted for at least twenty five years'. But I do not understand how there can be such an object.[1]

---

[1] And, of course, even if there can be, it could hardly count as an example of something whose identity was indeterminate because its (temporal) boundaries were *fuzzy* – but this is the conception which is the announced target of Evans' criticism.

My difficulty is not that I do not understand how an object can determinately satisfy a disjunctive predicate while determinately satisfying neither disjunct. I have no difficulty with the idea of an object (a 'borderline case') which is determinately either orange or red, but neither determinately orange nor determinately red. Similarly, I have no difficulty with the idea of a person who is determinately either a child or a young man, but neither determinately a child nor determinately a young man. But what we have been led to by applying Broome's conception of indefinite identity to the example of the club is different from this; it is more akin to the idea of a person who is determinately either a child or a senior citizen, but is neither determinately a child, nor determinately a senior citizen – and *is* determinately neither a young man nor a middle aged man.[2] And this is an idea I do find it hard to understand.

Fortunately, however, there is no need to struggle to understand Broome's example in terms of his own conception of indefinite identity. For we can give a perfectly intelligible account of the indeterminateness in truth-value of the statement 'the earlier club is the later club' without doing so. The source of the indeterminateness is simply that the terms 'the earlier club' and 'the later club' do not have determinate denotations. Let us fix on the former term. Because our notion of 'a club' is imprecise, there are, I suggest, two candidates for the denotation of this term, each of which would count as a club on some but not all acceptable sharpenings of the imprecise expression 'club'. One of these candidates is an entity which lasts a mere five years, and ceases to exist when the event which may or may not signal the demise of the earlier club occurs. Another is an entity which exists throughout those five years, and also exists throughout the twenty five year period after the members of the original group get together again at the end of the year of inactivity. Since one of these candidates lasts a mere five years, and the other at least twenty five (in fact, at least thirty) years, it is not puzzling that the statement 'the earlier club lasted either at most five years or at least twenty five years' is determinately true, while neither of the statements 'the earlier club lasted at most five years' and 'the earlier club lasted at least twenty five years' is determinately true.

The one difficulty which there might be about accepting this account of the example, I think, is the difficulty involved in believing that there *are* the two candidates for the denotation of 'the earlier club' whose existence it entails. In my earlier paper I discussed this difficulty in connection with an example of an (arguably) indeterminate statement of diachronic *personal* identity, viz. 'Brownson is Brown'. There I pointed out that the existence of rival candidates for the denotation of 'Brown' was only a difficulty if one did not accept a conception of continuants as possessors of temporal parts, for if one *did* accept that conception it was a total triviality. Exactly the same point applies in connection with the present example. Of course, many people *will* be reluctant to accept a conception of continuants as four-dimensional possessors of temporal parts, and for them the difficulty in believing in the existence of both of the rival candidates is, I think, in either example, a real one (in the earlier paper I suggested that this was an argument in favour of the four-dimensional conception); but I submit that the difficulty involved in believing in what Broome's notion of indefinite identity leads to when applied to these examples is far greater.

---

[2] Still more akin, of course, if the idea of a person who *at death* is determinately either a child or a senior citizen, but is neither determinately a child nor determinately a senior citizen and *is* determinately neither a young man nor a middle-aged man.

University of Birmingham,
P.O. Box 363, Birmingham B15 2TT

© H. W. NOONAN 1984

ON THE TRACKS OF THE SCEPTIC

By JONATHAN DANCY

ROBERT Nozick's recent *Philosophical Explanations* (Harvard and Oxford University Presses, 1981, pp. 167–288) offers what we might call a Conditional Theory of Knowledge. On this theory, 'a knows that p' is equivalent to the combination of

(1) p is true
(2) a believes that p
(3) If p were not true, a would not believe that p
(4) If, in changed circumstances, p were still true, a would still believe that p.

I agree with Crispin Wright that this theory succeeds in illuminating an aspect of our ordinary concept of knowledge (see Crispin Wright, 'Keeping Track of Nozick', ANALYSIS 43.3, June 1983, pp. 134–40). My interest, like Wright's, lies in Nozick's success in dealing with the sceptic. In this paper I wish to argue that Nozick is unsuccessful. My arguments are additional to Wright's, and do not depend on his. They also have an effect on Nozick's claim that his theory is externalist and hence escapes internalist objections (cf. David Gordon, 'Knowledge, Reliable Methods and Nozick', ANALYSIS 44.1, January 1984, pp. 30–3).

As a preliminary, I need to draw to attention to two ways in which Nozick's account of subjunctive conditionals such as (3) and

(4) differs from that of David Lewis. The Nozick version is that a subjunctive conditional $p \mathrel{\Box\!\!\rightarrow} q$ is true iff $p \supset q$ holds within a range of possible worlds reasonably close to the actual world. The first difference between the two accounts is that for Nozick, $p \mathrel{\Box\!\!\rightarrow} q$ is never true just because p and q are in fact both true, true in the actual world. Lewis takes it that $(p \,\&\, q) \rightarrow (p \mathrel{\Box\!\!\rightarrow} q)$, and relies (implausibly in my opinion) on Grice's theory of conversational implicatures to escape the counter-intuitive consequences (D. Lewis, *Counterfactuals*, Blackwell, 1973, p. 28). Nozick's account is different, as it needs to be; for otherwise (4) would be redundant, given (1) and (2). This difference is signalled by the phrase 'in changed circumstances' in (4). The second difference is that, for Nozick, the contradictory of a subjunctive conditional of the form 'If p were true, q would be true' will not normally be 'If p were true, q might still be false'. Whether these two sentences necessarily differ in truth value depends on the range of worlds we take ourselves to be considering. To find a conditional which is a genuine contradictory of a would-conditional, we have to find some means of ensuring that both conditionals are concerned with the same range of reasonably close possible worlds. For it is only within such a range that the truth value of $p \supset q$ is relevant to the truth or falsehood of $p \mathrel{\Box\!\!\rightarrow} q$. Since it is sufficient for the truth of the might-conditional merely that there be some world, maybe very distant, in which $p \supset q$ fails, the might-conditional will not normally be the contradictory of the would-conditional, which concerns itself only with closer worlds. To restrict our might-conditionals to the same group of close worlds as are relevant to the truth of the would-conditional, we need to use a new and perhaps unfamiliar connective 'might easily be'. Thus the contradictory of

(a) If p were true, q would be true

is

(b) If p were true, q might still easily be false.

In the present context the immediate relevance of this is that it enables us both to understand the origin of and to express the error in the classic sceptical argument that you cannot know that p if your belief that p might be false. This argument attempts to set a might-conditional against a would-conditional; Nozick's account shows why this cannot be expected to be successful. It would only be successful if or to the extent that we can show that your belief that p might *easily* be false, which is a much harder matter; the mere possibility of falsehood is irrelevant.

We can now turn to the central sceptical argument, which concerns brains in vats. You do not know that you are not a brain in a vat, being fed your current experiences by some ingenious technician/scientist (benevolent or malevolent according to taste). For if you were such a brain, then, provided that the scientist is successful, nothing in your experience could possibly reveal that you were; for your experience is *ex hypothesi* identical with that of something which is not a brain in a vat. Since you have only your own experience to appeal to, and that experience is the same in either situation, nothing can reveal to you which situation is the actual one. So you don't know that you are not a brain in a vat.

Is it possible, however, that though you do not know that you are not a brain in a vat you still know many other things, perhaps more important? Unfortunately, if you do not know this there is not much else that you can know either, it seems. Suppose that you claim to know that you are sitting reading ANALYSIS. You presumably also know that if you are sitting reading, you are not a brain in a vat. We can surely conclude that if you know that you are sitting reading, you know that you are not a brain in a vat, and hence by simple *modus tollens* that since you don't know that you are not a brain in a vat (agreed above) you don't know that you are sitting reading.

The principle of closure on which this argument relies can be formalized:

$$PC^k\text{: } (Kap \,\&\, Ka(p \Rightarrow q)) \rightarrow Kaq$$

This principle, given $\sim Kaq$ and $Ka(p \Rightarrow q)$ as in the above instance, allows us to infer $\sim Kap$. The principle seems therefore to show, more generally, that since you don't know that you are not a brain in a vat you cannot know any proposition p of which you know that if p were true, you would not be a brain in a vat. And there are similar slightly different arguments, for instance Descartes' version which takes q = 'you are not dreaming' and argues that since you don't know that you are not dreaming you don't know any proposition p such that if p were true you would not be dreaming.

What does Nozick have to say about this argument? He accepts its first move, that you do not know you are not a brain in a vat. Indeed, this thesis is supported by his Conditional Theory. For a necessary condition for such knowledge, (3), fails; it fails because if you were a brain in a vat (being fed your current experiences) you would still believe that you are not a brain in a vat. However, you do know that you are sitting reading, and that if you are sitting reading you are not a brain in a vat (all four conditions are satisfied in both these cases). This contradicts PC^k, but the Conditional Theory shows directly that PC^k is false. On Nozick's account of subjunctive conditionals, it is obvious that there will be some instances in which close worlds are so disposed that Kap and $Ka(p \Rightarrow q)$ are true, but in which truth conditions for Kaq concern not those worlds alone but more distant ones. Wherever this is true PC^k fails. It fails particularly blatantly where q concerns far-fetched suggestions such as your being a brain in a vat. The worlds relevant to the truth of Kaq lie vastly beyond the worlds relevant to the truth of Kap and Ka($p \Rightarrow q$). Therefore it is possible for Kap & Ka($p \Rightarrow q$) to be true and Kaq false.

Crispin Wright complains about the account of truth conditions for subjunctive conditionals, on the broad sweep of which Nozick is relying. I want to argue that even if we accept that account, Nozick is still unsuccessful in defusing the sceptic. Nozick supposes that all sceptical arguments rely on PC^k. But apart from the antecedent implausibility of this suggestion it seems that there is one sceptical argument at least which he cannot reject this way.

This is the argument by which it was originally established that you do not know you are not a brain in a vat. I take this argument to depend upon an epistemological version of the principle of Universalizability familiar in ethical theory. The principle holds that a judgement in one case commits one to making the same judgement of any relevantly similar case. Relevant similarity is defined in terms of the reasons for the judgement. And however broadly we cast the net, it will be at least true that in two cases that are completely indistinguishable to you, you must make the same judgements. The argument drives us to say that if you want to claim that you know now, you must be able to point to some recognizable difference between this case and other similar cases in which you (now) agree that you did not know. There is stress here on the word 'recognizable'. It excludes some properties, e.g. the property of truth, just as in ethics goodness is not held to be recognizable (which is why the principle of universalizability is not trivial).

You do not know you are not a brain in a vat because there is no difference which you can point to between your present situation and the imaginably similar situation in which you are a brain in a vat. In the latter situation you don't know that you are not a brain in a vat, trivially, and hence in the former you don't either.

Let us call this argument the 'argument from error' (though it works from possible as well as actual error). If it is on this sort of argument that we rely in order to show that we do not know that we are not brains in vats, various consequences follow. First, the argument from error is far more powerful than the one which relies on PC^k; the latter only casts doubts on claims to know propositions which entail that the knower is not a brain in a vat (or dreaming or something like that), but the former can be used to show that you do not know any proposition if you could not tell the difference between your present situation and an imaginably similar one in which the proposition is in fact false. (This is the thrust of the point that, generally, for all you know you might be wrong.)

Second, there is something right about a position which holds that for the very same sort of reason that you don't know you are not a brain in a vat, you don't know, e.g. that *The Times* will be published tomorrow, nor whether you are sitting reading. Surely it is implausible to suppose, taking Russell's example, that I do know what I did yesterday but I don't know that the world did not begin five minutes ago replete with archaeological and other traces. Nozick wants to show us that we can suppose this; but I am not sure that I really want to (or ought to want to).

Third, the argument from error cannot be attacked for relying on PC^k. This is not because PC^k is valid, particularly. Even if it were valid, I can see no reason for supposing that the argument relies upon it. Nozick anyway could not hold that it does, because his position relies on an independent proof that you don't know you are not a brain in a vat. Without that proof, we would hold it against the Conditional Theory that it can't even show that we don't know that we are not brains in a vat. With it, Nozick can and does find independent support for his theory; he wants to say that the theory gets things right here. But if the argument from error does rely on PC^k, which on his theory is invalid, he lacks that independent support and with it a considerable degree of plausibility. And I do not know of any other argument, not relying on PC^k, which Nozick could use to show that we do not know we are not brains in a vat without showing that we don't know most other things either.

Fourth, the argument cannot be rejected on the similar grounds that it argues directly from the fact that you might be wrong to the conclusion that you don't know. Nozick might try this line, holding that the contradictory of the relevant would-conditionals are not the true might-conditionals but the false might-easily conditionals. But, first, he needs the argument and cannot reject it; and, second, the argument does not proceed directly in this way, which would be fallacious, but indirectly via the Principle of Universalizability.

Finally, there is the distinction between externalism and internalism in epistemology. This distinction, as revealed in the literature, is extremely unclear; and I cannot here do anything to improve matters. It would normally be agreed, however, that Nozick's theory is externalist because he does not add to (1)–(4) any of the following:

(5) $\mathrm{Ka}((p \mathrel{\Box\!\!\rightarrow} \mathrm{Bap})\ \&\ ({\sim}p \mathrel{\Box\!\!\rightarrow} {\sim}\mathrm{Bap}))$
(6) $\mathrm{Ba}((p \mathrel{\Box\!\!\rightarrow} \mathrm{Bap})\ \&\ ({\sim}p \mathrel{\Box\!\!\rightarrow} {\sim}\mathrm{Bap}))$
(7) a has some cognitive grasp of the reason why $p \mathrel{\Box\!\!\rightarrow} \mathrm{Bap}$ and ${\sim}p \mathrel{\Box\!\!\rightarrow} {\sim}\mathrm{Bap}$

(For (7) as a version of internalism cf. Bonjour, 'Externalist Theories of Empirical Knowledge', in *Midwest Studies in Philosophy V*, eds P. French, T. Uehling and H. Wettstein, University of Minnesota Press, Minneapolis, 1980, p. 55). Nozick might wish to say that his conception of knowledge is externalist, while the second sceptical argument is crucially internalist. If externalism is acceptable, he can therefore dismiss our argument as irrelevant. This sort of move is made in defence of Nozick by David Gordon (op. cit.). But it is not available, for two reasons. First, again, Nozick relies on the argument, internalist or not, to provide the independent support for something which would otherwise look like a counter-intuitive consequence of his theory. Second, in taking the argument to be necessary Nozick tacitly accepts that his theory is not as purely externalist as it might initially appear, and hence he cannot claim immunity from an attack merely on the grounds that it is internalist.

I have argued that even if PC^k is invalid, there remains a different sceptical argument which establishes that we do not know that we are not brains in vats. There must be such an argument if there is to be any point in disproving PC^k, for PC^k uses that argument's conclusion as a lever for contraposition. But if there is such an argument, disproving PC^k cannot be a complete reply to the sceptic. In fact the relevant argument is both more powerful than the one Nozick considers and not vulnerable to his criticisms. The spectre of scepticism cannot be removed merely by disproving one principle of inference.[1]

University of Keele,
Keele, Staffordshire ST5 5BG

© JONATHAN DANCY 1984

---

[1] My thanks to David Bakhurst, Jennifer Hornsby, Charles Swann, Richard Swinburne and Clive Borst for helpful comments.

SPECIAL RELATIVITY AND THE LOCATION OF MENTAL EVENTS

By DAVID GORDON

MICHAEL Lockwood has proposed an ingenious method according to which mental events can, at least in some instances, be located in space (ANALYSIS 44.1, January 1984). If one grants that mental events occur in time, then, he holds, the special theory of relativity applies to them. It is this premise that underlies Lockwood's method for locating mental events. Unfortunately, this premise makes Lockwood's method useless to counter the sort of opposition to the identity theory he wishes to refute.

The opponent of the identity theory, as Lockwood presents him, denies that, apart from the identity theory, there is reason to assume that mental events are located in space, though he admits they occur in time. It seems to me that, in making this claim, the opponent of the identity theory wishes to question whether mental events are physical events at all; it is surely implausible to take him to be admitting that mental events *are* physical, only not spatially located. If he is right, and mental events are not physical, then

Lockwood's application of special relativity cannot be used to locate them, as the theory applies only to physical events. So far as I have been able to determine, this point is not controversial in standard expositions of the special theory. James L. Anderson, in *Principles of Relativity Physics* (Academic Press, New York, 1967), pp. 137–8, observed: 'It was Einstein's analysis of space-time measurements more than anything else that led to the acceptance of the relativity theory. It was clear from the start that Einstein considered these measurements in terms of comparisons between different physical systems. ... Thus we should not demand that what the physicist calls a space measurement or a time measurement bear much relation to our psychologically conditioned senses of these concepts. All that we must require is that such measurements should be realizable by real physical systems.' Similarly, C. Møller, *The Theory of Relativity* (Clarendon Press, Oxford, 1952), p. 4, states: 'Now the basic assumption of the special theory of relativity is that the *special principle of relativity is valid for all physical laws.*' (Emphasis in original: a footnote exempts the laws of gravitation from the sphere of the special principle.) This fact is hardly surprising; theories of physics by definition deal only with physical events. It is only if one first assumes that mental events are physical that one can use Lockwood's method to locate them in space. Since this is just what the critic of the identity theory challenges, Lockwood's refutation of him assumes the essential point at issue.

Lockwood might in reply appeal to the fact that, as he notes, mental events can both cause physical events and be caused by them; but unless one grants that only a physical event can be causally related to physical events, this will not help his argument.

6238 Orange Street,
Los Angeles, California 90048, U.S.A.

© DAVID GORDON 1984

REPLY TO GORDON

By MICHAEL LOCKWOOD

THE issue, as between Gordon and myself, is not, I take it, the purely exegetical one of whether the theory of relativity was ever intended to apply to mental events. Rather, it is the philosophical issue of whether mental events can coherently be excluded from the range of application of the theory. My contention is that that they cannot be so excluded. In the first place, the theory of relatively, in its mature form at least, is not just a theory about physical events. It is also a theory about the space and time in which such events occur. According to the theory, space and time have no reality independent of each other, any more than do the three dimensions of space.

Thus, I do not need to argue that mental events are physical, in order to justify the application to them of the theory of relativity. (Nor would I wish to appeal to any such principle as that 'only a physical event can be causally related to physical events'; that, I agree, would be utterly question-begging.) No, all I need argue is that mental events are in time, that they belong to the very same temporal order as do physical events. Given that the theory of relativity is, *inter alia*, a theory about the nature of time, it will then follow inescapably that the theory does apply to mental events, whatever physicists may or may not say on the matter. Specifically, it will follow that mental events, by virtue of being in time, are also in space-time. And from that, as I argued in my original article, it will in turn follow that they are in space.

How might this conclusion be resisted? Only, it seems to me, by denying that mental and physical events belong to a common temporal order. To be sure, it is a part of our common sense view of reality that mental and physical events have a shared chronology. But then it is independently obvious that the theory of relativity requires us to revise some, at least, of our common sense preconceptions: common sense, in this area, is thus far from sacrosanct. What I would appeal to, in defence of the common sense view, is the following principle: that two events can be causally related only if they are also temporally related. I am inclined to think that this is a conceptual truth; but at any event, the prospect of having to deny it would, I am confident, strike most people as singularly unappealing. Given the scarcely deniable fact that mental events can both cause and be caused by physical ones, it then follows, by the above principle, that they do, after all, belong to the same temporal order.

A weakness of Gordon's argument, as it stands, is that, while he claims that mental events may be exempted from the strictures of the theory of relativity, he nowhere indicates what sort of a picture of time, space and psycho-physical causation he envisages as following from this. If he, or anyone else, really thinks that it is logically consistent to hold relativity true of physical events but inapplicable to mental ones, without denying psycho-physical causation, let him provide a model, a properly articulated description, of what he has in mind. The onus will then be on me to demonstrate that what has been offered is either internally inconsistent, incompatible with the theory of relativity, or else contradicts some incontrovertible fact. The ball is in his court.

Department for External Studies,
Oxford University,
Wellington Square, Oxford OX1 2JA

© Michael Lockwood 1984

PREFERENCE AND TRANSITIVITY

By RICHARD LEE

I WILL argue that preferences which violate transitivity are inconsistent. We are all familiar with preferences. You may prefer attending a movie tonight to writing a philosophy paper. You may prefer that federal funds to social programmes not be cut. But what is a preference?

As an approximation we may say that to prefer A over B is to be disposed or inclined to choose A over B. I do not mean to be insisting on a strict behaviourist definition of preference, which might read: one prefers A to B just in case one in fact would choose A instead of B when given the option; just in case, that is to say, one would never choose B when one could choose A instead. (One might choose some third thing instead of either of them.) Rather, a preference is a state of mind which plays a role in one's psychological makeup in general or at a specific time. One has a preference when one is set to choose one thing over another. This does not guarantee that one in fact would make the specified choice when the options were actually presented. For one thing a person may act absent-mindedly. One can absent-mindedly act against one's beliefs as well as against one's preferences. It may be that I do prefer listening to the radio to watching television, but when the choice came up I was distracted. Thus the fact that someone has a preference is not a simple fact about that person of the form 'A would X if C' (although it is a fact about the person nonetheless).

I have said that a preference can be understood as a disposition to choose. Now it is obvious that many preferences which we have concern situations in which there is no question of *choice* on our part. I might prefer tomorrow to be sunny rather than rainy. I of course have no say in the matter. Still the preference relates to a *possible* choice situation. It is merely factually or contingently impossible, so to speak, for someone to choose what tomorrow's weather is to be. It is not metaphysically impossible. We understand what it would mean to make such a choice even though we do not readily understand how such matters could be put into our hands. Similarly it is now impossible for us to affect the past. Yet we may prefer that the war in Vietnam never have been fought. We are not now in a position to make such a choice, but we can understand what it would mean for someone at some time to have been in such a choice situation.

What are the objects of preferences? On the face of it they are a motley sort. One can have preferences among drinks, candidates for office, colours, movies, and practically anything else. I propose that all preference can be understood in terms of preference among states of affairs or possible circumstances. A preference for bourbon, for example, may be a general preference that one drink bourbon instead of drinking scotch. The objects of preference are thus nearly enough the same as the objects of belief. But preferences, unlike beliefs, have multiple objects. They are bipolar. A preference is for one thing over another. I will symbolize this by 'P(A, B)'.

Often preferences are expressed in conditional form. I prefer drinking coffee to drinking tea when I am eating chocolate. This does not imply that I always prefer coffee to tea. We may say that someone prefers A to B given C when that person is disposed to choose A over B when C. I will symbolize such a conditional preference by 'P(A, B/C)'. This type of conditional preference must be distinguished from something else which might also naturally be called 'conditional' preference. Let us suppose that when I am depressed I prefer playing my guitar to reading philosophy. What this means is that under certain conditions I am disposed to behave in certain ways. Here it is the case that, as a matter of fact, if I am depressed (whether I am aware of it or not) I will have a certain preference; I will be disposed to make a certain choice. But this is different from the notion of conditional preference I am concerned with. In these cases, of which the coffee/tea/chocolate preference is an example, I am *now* disposed in a certain way, namely to pass up tea for coffee when I eat chocolate. It might be that were I actually to eat some chocolate I would not *then* be disposed in the same way. So in one sense of 'conditional preference' when one has such a preference there is a present disposition, but a qualified one as in the coffee/tea/chocolate case. I will call this 'internal conditionalization' of a preference. External conditionalization is when whether one has a certain preference depends upon what the situation is, as in the depression example. An analogous distinction can also be drawn in the realm of belief. The analogue to an internally conditional preference is a belief in an if-then proposition. External conditionalization is markedly different. It may be true of someone that if she were kidnapped she would come to believe her kidnappers to be wonderful people. This is quite different from her now believing that if she were kidnapped her kidnappers would be wonderful people. The former is an example of external conditionalization, while the latter is a case of internal conditionalization.

A conditional preference in the external sense is not, as such, a preference. It is not itself a state of mind. Rather, it is a proclivity for coming to have a preference under certain circumstances. It may be the case that as a matter of fact if Chris gets socked in the jaw he will prefer to strike back. But that is not what he now prefers — even conditionally. He prefers that if he be struck he turn the other cheek. Thus the existence of external conditional preferences requires no revision in what was said so far in my account of preference.

Conditional preference in the internal sense is on different ground, however. If I have a conditional preference for coffee over tea when I am eating chocolate, the object of this preference is not exhausted by the ordered pair of states of affairs ⟨my drinking coffee and my drinking tea⟩. We have a new term which is not a polar term. It is what may be called a *conditioning term*. (The conditioning term in P(A, B/C) is C.)

Conditional preferences of the internal sort can be formulated in terms of nonconditional preferences having two terms. Consider the chocolate example above. I have a preference for drinking coffee instead of drinking tea when I eat chocolate. But this same preference can be expressed without mention of a separate conditioning term. The preference is for *drinking coffee while eating chocolate* over *drinking tea while eating chocolate*. (Obviously whether I prefer to eat chocolate at all is an independent matter.) Thus P(A, B/C) is to be understood as equivalent to P(A & C, B & C). The polar terms are slightly more complex than before but they continue to be possible states of affairs (as long as C is logically consistent with A and with B).

I suggested earlier that to prefer A to B given C is roughly to be disposed to choose A over B when C. The reduction of conditional preferences to nonconditional preferences can be defended in terms of this account. If C is given, then the options under consideration do not include A and B *simpliciter*, but (A & C) and (B & C). To prefer A to B given C is precisely to prefer (A & C) to (B & C). If I am already eating chocolate, then a choice of coffee over tea is a choice of coffee with chocolate over tea with chocolate.

As is the case with beliefs, not all sets of preferences are consistent. A preference and its converse, for example, are jointly inconsistent. I use the terms 'consistent' and 'inconsistent' to denote synchronic rather than diachronic relationships. If Frank is fickle and frequently changes his preferences among women to sleep with for no apparent justificatory reason, he is being *inconstant*, but not therefore being inconsistent in the sense I am using that term. I will say that *a set of preferences, R, is inconsistent if there is some possible set of options (states of affairs), O, such that the choice of any option in O would contravene some preference in R. A choice of B from an option set containing A and B would contravene the preference P(A, B).* The preferences P(A, B) and P(B, A) are together inconsistent, then, because any choice from the option set {A, B} would contravene one of the preferences.

I submit that this notion of inconsistency is a natural one. There are plenty of situations in which a particular choice is not demanded by a set of preferences. But if there are circumstances in which no matter what one chooses one is going to go against some preference one holds, then there is something amiss with one's preferences: one cannot really be disposed to choose strictly (i.e. in all circumstances) in accordance with all the preferences. (There is nothing 'amiss' in this sense just because one cannot always get what one prefers. That is simply a fact of reality.)

Given this notion of inconsistency it is easy to show that atransitive preferences (as I will call preferences which violate transitivity) are inconsistent. One cannot prefer A to B and B to C while preferring C to A without being guilty of an inconsistency. To prove this we need only consider the option set consisting of A, B, and C. Choice of any of these would be contrary to, i.e. would contravene, one of the preferences in the set.

In *The Logic of Preference* (section 9, p. 22) Georg von Wright is troubled by the following case:

> Suppose a person is offered, 'on a tray', a choice between an orange and a banana. He chooses the orange. He is then offered a choice between a banana and an apple. He chooses the banana. Finally he has to face a choice between an orange and an apple. Is he then 'by the laws of logic' bound to choose the orange? Does he 'contradict' himself, in thought or action, if he chooses the apple?

(von Wright thinks transitivity holds, at least for the sort of preference he is interested in, but the passage from which this quote is taken makes it clear that he is nevertheless troubled by this example.) Of course it is possible for such a person to choose the apple on the final round. It is also fairly obvious that he would by no means be contradicting himself — but neither is he violating the transitivity of preference. His preferences could have changed. Yet even if we assume the person's basic preferences to remain constant, there is not a violation here. The *situation* has changed. By the time the person is offered the choice between the orange and the apple he has, we are told, already chosen an orange on a previous occasion. He may quite rationally prefer having an apple *and* an orange to having two oranges and no apple. It would even be rational to choose an orange over an apple and then choose an apple over an orange the next time the tray came his way. He may simply prefer having a variety of fruits. (One preference I suggest von Wright's chooser may have is this: P(I have an apple, I have another orange/I already have an orange). This is consistent with the preferences P(I have an orange, I have a banana/I have no orange) and P(I have a banana, I have an apple/I have no orange).)

It is only when the conditioning term is constant that the preferences can legitimately be called atransitive. Atransitive preferences are those which violate the principle of transitivity: if P(A, B) and P(B, C), then P(A, C). Given the remarks about conditional preferences above, it can be seen that the law applied to conditional preferences would be: if P(A, B/D) and P(B, C/D), then P(A, C/D).

Still some philosophers believe that atransitive preferences are not only possible, but may be perfectly rational. In a recent article in ANALYSIS ('Rationality and Intransitive Preferences', ANALYSIS 40.3, June 1980) R. I. G. Hughes argues that in some cases atransitive preferences may be held consciously and rationally. Hughes offers as an example the preferences of an imaginary voter:

> [Given] the choice between Aley and Baley, he would vote for Aley on the basis of his political programme, since although Baley is the more honest, the difference between them is small enough for that consideration to be unimportant. Similarly he would vote for Baley over Caley. However, faced with the choice between Aley and Caley, he would find the difference in corruptibility too great to ignore; this difference would override his purely political preferences and he would vote for Caley. (p. 133)

A person may of course be prepared to vote in this way and may have preferences which reflect this and be no less rational for it. But just what are the preferences here? They are not simply P(A, B), P(B, C) and P(C, A). They are, it seems, rather P(A, B/the choice is between A and B), P(B, C/the choice is between B and C), P(C, A/the choice is between A and C). But these preferences do not violate transitivity since the conditioning term is not constant. Similarly there could be moving objects such that A is moving faster than B given reference frame 1, B is moving faster than C given reference frame 2, and C is moving faster than A in reference frame 3. Yet 'is moving faster than' is still a transitive relation. That the preferences of Hughes' voter, as I have described them, are not atransitive is even more obvious if the conditioning terms are absorbed into the polar terms in the way previously explained.

Someone may respond that the preferences of Hughes' voter can be described with a fixed conditioning term, namely 'the choice is between two candidates'. This may seem to circumvent the solution above. On my account these preferences with a fixed conditioning term *would* violate transitivity – and they would be inconsistent. Letting T be the conditioning term we simply consider the option set {A & T, B & T, C & T}. But perhaps this simply shows that these are not the preferences of the voter in the situation Hughes sketched. There the voter is asked in effect 'if these were the two candidates, who would you choose?'. But this identifies the preferences as having distinct conditioning terms depending upon which candidates are being compared. (This argument may appear to be beside the point, but Hughes' arguments for the reasonableness of his voter's preferences (later on p. 133) depend upon which two candidates in any case are presumed to appear in the election. His voter considers whether 'there is a markedly more scrupulous candidate available'). If the conditioning term is held constant, the question addressed is somewhat different: 'Knowing that the election will be between two candidates, but not knowing which two these will be, who would you prefer to win the election?' The question can then be made more restrictive by adding 'of Aley and Baley?', 'of Baley and Caley?', etc. It is precisely these sorts of questions which Hughes' voter is not prepared to answer. (I think Hughes could admit this point. See p. 134 of his article. He would claim that this is a disguised way of asking for a rank ordering of the candidates, and *that* his voter is not prepared to offer.) I do not say that Hughes' voter is irrational for not being able to answer these questions, but if he cannot answer them, his preferences do not have a constant conditioning term and, while not irrational, the preferences are thus not atransitive either.

University of Arkansas,
127 Communications Center,
Fayetteville, Arkansas 72701, U.S.A.

© RICHARD LEE 1984

PUTNAM'S MODEL-THEORETIC ARGUMENT AGAINST METAPHYSICAL REALISM

By ANTHONY L. BRUECKNER

IN this note, I wish to state an interpretation of Hilary Putnam's 'model-theoretic argument' against 'metaphysical realism' ([1], [2], [3]) and then to indicate the force of a response to the argument which Putnam himself anticipates but underestimates.

A *metaphysical realist*, as understood by Putnam, espouses a correspondence theory of truth according to which the truth of a theory (the truth of the members of a set of sentences) depends upon (1) *the intended interpretation of the theory* (the intended model) and (2) the way the world is. For example, a theory's sentence 'The cat is on the mat' is true just in case the object denoted by 'the cat' (according to the theory's intended interpretation) bears the relation denoted by 'is on' (according to the intended interpretation) to the object denoted by 'the mat' (according to the intended interpretation). Putnam wants to argue that

> (P) If the metaphysical realist conception of truth is correct, then truth is radically indeterminate.

But since truth is not radically indeterminate, it follows that the metaphysical realist conception of truth is incorrect. Instead, we must adopt an 'epistemic' conception of truth according to which a proposition is true just in case it would form a part of the ideal theory. On such a conception, it makes no sense to suppose that the best theory about the world which humans are able to come up with might in fact be false (owing to its failure to correspond to the way the world is). In this paper, I will focus upon whether (P) can be rejected in the face of Putnam's remarks in support of it. I will not consider the following two ways of blocking Putnam's argument as just sketched. (A) Accepting that truth is radically indeterminate in the sense that the intended interpretation for a theory *cannot* be specified. (B) Denying that an 'epistemic' theory of truth is the only alternative to a correspondence theory.

How does Putnam attempt to establish (P)? Suppose that I say that on the metaphysical realist conception of truth as correspondence to the world, it is in fact quite determinate whether my sentence 'The cat is on the mat' is true. This is because it is quite determinate whether the world is the way that the sentence says it is according to the intended interpretation of the sentence. However, Putnam says, there are many interpretations which make the sentence true and many which make it false. Unless there is some way of specifying which interpretation is the intended one, it makes no sense to say that the sentence is determinately true (or determinately false) *relative to its intended interpretation*. Putnam focuses upon 'unintended interpretations' which make a sentence or theory *true*. He focuses upon unintended *models* in order to make the point that if a metaphysical realist maintains that our ideal theory might be *false*, then Putnam can challenge him somehow to rule out an interpretation which makes the theory *true* (e.g., a model in the natural numbers). However, I think that Putnam's point can be stated most generally as

> (Q) If the metaphysical realist cannot specify which interpretation of a theory T is the intended one, then the metaphysical realist's conception of truth leaves it radically indeterminate as to whether T is true.

If, e.g., I cannot rule it out that a theory containing 'The cat is on the mat' has its intended interpretation in the natural numbers, then I leave the truth of the theory radically indeterminate. And if the theory is first order and is consistent, it will have countless models in the natural numbers (as well as countless interpretations in the natural numbers which make some or all of its member sentences false).

Consider the theory T_1 containing just these three sentences which, we will suppose, I assent to:

(1) The flea is on the cat.
(2) The cat is on the mat.
(3) The mat is on the floor.

The following interpretation obviously makes (1)–(3) all true: 'the flea' refers to the number one, 'the cat' refers to the number two, 'the mat' refers to the number three, 'the floor' refers to the number four, and '— is on . . .' refers to the relation — *is numerically less than* . . . . An interpretation just like the foregoing except that '— is on . . .' refers to the relation — *is numerically greater than* . . . obviously makes (1)–(3) all false. But why can't I specify the *intended* interpretation of T_1 as follows:

(4) 'the flea' refers to Fred
(5) 'the cat' refers to Puss
(6) 'the mat' refers to my living room mat
(7) 'the floor' refers to my living room floor
(8) '— is on ...' refers to the relation — *is spatially on top of...*

Given this specification, and given the spatial relations obtaining among Fred the flea, Puss the cat, my living room mat, and my living room floor, one would be able to claim that (1)-(3) are as determinately true as one could desire.

Putnam's response is that (4)-(8) are themselves just more bits of language, more theory. The theory T'_1 consisting of (1)-(8) has countless interpretations, for example, in the natural numbers which make T'_1 true, and T'_1 also has countless interpretations in the natural numbers which make some or all of the members of T'_1 false. Putnam will put the point in the following way: on the metaphysical realist conception of 'intended interpretation', the use of a language (or theory) does not fix the interpretation of that language (or of that theory). (See [3], p. 466 and p. 482.) Since language does not fix its own interpretation, using more language (e.g., using sentences (4)-(8)) in order to fix the interpretation of the bits of language (1)-(3) is fruitless. The second bits of language are no better off than the first, in the sense that if there is a problem as to the interpretation of the first bits, there will be just as much of a problem as to the interpretation of the second bits. (See [2], pp. 126-7, for Putnam's graphic way of making this point.)

Now these considerations might seem to establish that

(S) If the metaphysical realist conception of truth is correct, then the metaphysical realist cannot specify which interpretation of a theory T is the intended one.

If we adopt the metaphysical realist conception of truth as relative to an intended interpretation, then we seem unable, e.g., to specify the non-mathematical 'intended' interpretation of our theory T_1: the attempted specification ((4)-(8)) seemed to fail for the reasons given in the preceding paragraph. The same reasoning applies to the attempt to specify the intended interpretation by appeal to what Putnam calls 'operational and theoretical constraints' ([1], pp. 29-32). That is, the same problems about interpreting T_1 will arise for the theory $T_1\#$ consisting of the union of T_1 and the set of sentences (apparently) expressing the operational and theoretical constraints, whatever they might be. Now if Putnam *has* established (S), then, given the plausibility of (Q), we can derive (P), which on my understanding is the key contention of Putnam's model-theoretic argument against realism.

The most that is established by the considerations of the last paragraph but one, though, is that the metaphysical realist's use of language in an attempt to specify a theory's intended interpretation does not *by itself bring it about* that a particular interpretation is the *intended* interpretation. The sense in which it is true that the use of (4)-(8) does not by itself 'fix' some interpretation as the intended interpretation of T_1 is the foregoing *metaphysical* sense. The mere *use* of the sentences comprising T_1 obviously does not fix their interpretation: does not make it the case that some interpretation is their intended interpretation. Similarly, the use of the sentences (4)-(8) does not by itself fix *their* interpretation. Hence the use of (4)-(8) does not fix the intended interpretation of any other sentences either (so does not fix the intended interpretation of T_1). Putnam also argues that some interpretation's being the intended interpretation cannot either be brought about – made to be the case – by a use of language *backed by intention*. Insofar as an intention is a representational state, the interpretation of its content is subject to whatever difficulties beset the interpretation of a sentence. But all this leaves open the possibility that some non-linguistic (non-representational) feature of reality, together with the use of language (possibly backed by intention), bring it about that a theory (stated using the language) has a particular interpretation as its intended interpretation. For example, it could be maintained that the causal connections between my utterances of (1)-(8), on the one hand, and extralinguistic reality, on the other, bring it about that a particular interpretation (call it I) is the intended interpretation of the theory T_1' comprising (1)-(8). It could be maintained, further, that relative to I, (1)-(8) are all true. If that were the case, then (4)-(8) *would* be a correct specification of the intended model of T_1. Then (S) above would be false (or at least would not yet have been supported), and Putnam's argument for (P) would be blocked. The model-theoretic argument would thereby be blocked.

The strategy, then, is to agree that the interpretation of a language is not fixed by its use. We nevertheless disagree that (S) follows. We maintain that the use of language together with non-linguistic facts about the world (e.g., causal relations between the world and the use of language) do fix the intended interpretation of the language. Unless Putnam can discredit the foregoing claim, he has given us no reason to suppose that every attempt at a linguistic specification of a theory's intended interpretation must fail. That is, unless he can discredit the claim in question, he has given us no reason to suppose that a specification of T_1's intended interpretation such as (4)-(8) is not a correct specification *given the intended interpretation of (4)-(8)*.

Putnam is aware of this sort of response to his model-theoretic argument (see [1], pp. 45-8). He considers that a causal theorist of reference such as Hartry Field might espouse a physicalistic definition of reference

(D) *x refers to y* iff *x bears R to y*

(in which R is a physicalistic relation 'definable in natural science vocabulary without using any semantical notions'). The claim, then, would be that (D) describes the way in which the intended interpretation of a theory such as T_1 is fixed by use of language together with non-linguistic features of reality. Further,

(D*) The fact described by (D), together with the use of (D), fixes the intended interpretation of (D) itself.

In particular, the term 'R' refers to a kind of causal relation by virtue of the existence of a causal relation of that kind between (a) instances of the relation in question and (b) uses of the term 'R'.

Is there a problem here? Putnam says that the claim of (D*) 'is not necessarily expressed by just *saying*' what is said in (D) ([1], p. 46). This may be true, but we can use (D*) itself to express the claim in question. The next question which Putnam raises is:'If ... [(D)] is true, ... understood [as describing the determinate unique relation between words and their referents], what *makes* it true? ... It seems as if the fact that R *is* reference must be a *metaphysically unexplainable* fact, a kind of primitive, surd metaphysical truth' ([1], p. 46). Would anyone take the following to be a serious objection to the traditional analysis of knowledge: 'The traditional epistemologist has produced nothing but the unexplainable, primitive, surd metaphysical "truth" that knowledge just *is* justified, true belief'. If this is not a reasonable objection in the case of the traditional account of knowledge (which of course has *other* problems), then why should it be a reasonable objection in the case of the causal account of reference (which may well have *other* problems)?

Putnam's remarks in expanding upon the current problem are often misleading. For example, he says that 'to me, believing that some correspondence intrinsically just *is* reference (... as an *ultimate* metaphysical fact) amounts to a magical theory of reference' ([1], p. 47). However, he introduces the pejorative term 'magical theory of reference' in the context of considering the example of a Martian who has had no causal contact with trees (and no contact with anyone who has had contact with trees) yet comes to form a mental image exactly similar to *my* mental image of a tree (as a result of seeing a paint blob exactly resembling a tree). To mistakenly suppose that the Martian's image is an image *of* a tree (refers to trees) would be to espouse a magical theory of reference. ([1], chapter 1.) But those who espouse (D) and (D*) would completely agree with Putnam's intuitions about the Martian. Causal theorists of reference hold that given the truth of (D), it obviously follows that the Martian cannot magically refer to entities with which it has had no causal contact.

Another misleading remark is that if (D) is true, then 'reference becomes what Locke called a "substantial form" (an entity which *intrinsically* belongs with a certain name)' ([1], p. 47). A similarly misleading remark is that the metaphysical realist's conception of truth as correspondence commits him to a belief in 'Self-Identifying Objects – objects that intrinsically correspond to one word or thought-sign rather than another' ([1], p. 51). But a realist who adopts (D) (or some other strategy on which appeal is made to facts about extra-linguistic reality) is not committed to the absurd views just described. It is not as if a causal theorist of reference, for example, must maintain that each name of English is somehow magically metaphysically paired with some referent, so that 'Hilary Putnam' *had* to end up referring to a certain person no matter what. No entity intrinsically belongs with any name, according to (D), since there is no necessity as to which entities bear which causal relations to which uses of which names. According to (D), no object *necessarily* bears its actual name, and the relation between a name and its referent is no more magical than any causal relation. It is worth noting that the foregoing defence of (D) is consistent with the view that names are rigid designators. It is consistent to hold that (1) there is a single individual which is such that for every counterfactual situation I might wish to speak about, when I use the name 'Plato' in speaking about the situation, the name denotes the individual (if it denotes at all), and (2) in some counterfactual situations, the individual in question does not bear the name 'Plato' in the language used in the situation.

As a parting shot, Putnam says 'Why could it not be a surd metaphysical fact that reference *is* the relation: x refers to y in at least one admissible model M?' ([1], p. 48.) An admissible model for a theory T is a model for a theory T# resulting from conjoining sentences expressing operational and theoretical constraints to the original theory T. So on this proposal, a name which figures in a theory will have as many referents as there are admissible models for the theory: a name, then, will have countless referents. Certainly it is much more plausible to suppose that (D) expresses a primitive metaphysical truth about reference than to suppose that this proposal does.

I conclude that if we understand Putnam's model-theoretic argument in the way that I have suggested, then he has no strong reply to the kind of objection developed here (following Putnam's own suggestion). Anyone who has read *Reason, Truth, and History* and the papers [2] and [3] will realize that it is not completely clear how the argument is supposed to proceed, so it will be open to Putnam to object that I have misunderstood him. Nevertheless, my presentation of his argument and discussion of the objection which I find compelling do have the virtue of showing this: the argument seems to force us to adopt some or other account of reference according to which extra-linguistic facts help fix reference. One might object that the argument does not even accomplish this much. In this note, though, I have mainly been concerned to establish that one can plausibly block Putnam's argument by holding some such view about reference.

Yale University, New Haven,
Connecticut 06520, U.S.A.

© ANTHONY L. BRUECKNER 1984

REFERENCES

[1] *Reason, Truth, and History* (New York: Cambridge University Press, 1981), chapters 2 and 3.

[2] 'Realism and Reason', in *Meaning and the Moral Sciences* (Boston: Routledge and Kegan Paul, 1978).

[3] 'Models and Reality', *Journal of Symbolic Logic*, volume 45, no. 3 (September, 1980).

PHENOMENALISM AND OBSERVATION CONDITIONS

By ERIC RUSSERT KRAEMER

IN his recent discussion, 'Chisholm's Objection to Phenomenalism' (ANALYSIS 42.1, January 1982) Michael J. Maloney criticizes R. M. Chisholm's classic attack on phenomenalism (in *Perceiving, A Philosophical Study*, Cornell, 1957; unchaperoned page references in the text will be to these two discussions). Maloney argues that Chisholm's objection does not succeed. Maloney's remarks, however, do not seem to offer any real comfort to the phenomenalist. In what follows I shall try to demonstrate why.

In the Appendix to *Perceiving* Chisholm urges that, contrary to what the phenomenalist holds, 'the familiar facts of "perceptual relativity" suggest that our ordinary thing statements do *not* entail any statement referring solely to appearances' (p. 190). For, Chisholm says, '...no synthetic thing statement P entails any appearance statement unless P is taken in conjunction with some *other* thing statement referring to observation conditions' (pp. 193-4). Maloney interprets Chisholm's argument as follows:

> (A) For any [Physical Object Statement] POS (P) alleged to entail a [Sense Datum Statement] SDS (S) there is an observation condition statement (Q) consistent with P such that P & Q entails ~S. Therefore it is not the case that any POS entails a SDS. (p. 25)

Maloney constructs two objections to (A). First, he claims, since 'Chisholm grants that the conjunction of P and some other thing statement Q′, referring to observation conditions, does entail S' (ibid.), the conjunctive physical object statement $P' = (P \& Q')$ does entail S, a SDS. Second, he asks us to consider another type of conjunctive statement, P″, which is formed by conjoining 'one POS (P) and the negations of any observation condition statements (Q) such that Q is consistent with P and the conjunction of P and Q entails ~S' (ibid.). Maloney asserts that '... no Chisholmian counterexamples can be constructed to defeat the claim that P″ does indeed entail S' (pp. 25–6).

These two objections, if correct, would show that the conclusion of (A) is false. It should be noted that Chisholm puts forward his argument with respect to '...our ordinary statements about physical things – such statements as "That is a tree" and "This thing is red"' (p. 190), not with respect to 'any POS'. Thus Maloney's two objections to (A) do nothing to show how the phenomenalist is to avoid what Chisholm takes the relativity of perception to demonstrate, namely what he later calls (in 'Theory of Knowledge in America', reprinted in *The Foundations of Knowing*, University of Minnesota Press, 1982, p. 180)

> a difficulty in principle for the attempt to translate thing statements into appearance statements: whenever we seem to reach such a translation we find that our translation still includes some set of thing statements that describe physical observation conditions and that have not been translated into statements about appearances.

If Maloney's objections were correct, then some conjunctive POS's would entail some SDS's; but this does not provide any clue as to how POS's in general are to be given a SDS analysis. In response, one might urge that the phenomenalist could take comfort from there being *some* recognized entailments between the class of POS's and the class of SDS's, and thus might be justified in thinking that a complete phenomenalistic analysis of POS's only awaited further effort. Given this response, Maloney's two cases need to be carefully examined. It is my contention that they do not support phenomenalism.

Both of Maloney's criticisms crucially depend upon the claim (*C*) that the conjunction of a POS (P) and some observation condition statement Q does entail a SDS. Maloney takes observation condition statements to '... refer to such things as the lighting conditions and the condition of someone's sense organs' (p. 25). But under what conditions is claim *C* correct? Consider the following example: [Example 1]

(P_1) The flower is yellow.
(Q_1) Blue lights are shining on the flower.
(S_1) Greenness will be sensed.

Does S_1 follow from $(P_1 \& Q_1)$? Clearly not, for the conjunction $(P_1 \& O_1)$ is compatible with further claims such as

(T_1) Red lights are shining on the flower.

And the conjunction, $((P_1 \& Q_1) \& T_1)$ does not entail S_1. So not every observation condition will do. Consider, however, Chisholm's example. In speaking of the relation between the sentence, 'This is red' (P) and the sentence 'Redness will be sensed' (R) Chisholm says: [Example 2]

> Taken in conjunction with certain *other* thing statements Q, referring to observation conditions, P does entail R. The following is such a statement Q:
>
> > This is perceived under normal conditions; and if this is red and is perceived under normal conditions, redness will be sensed. (Q) (192)

This example of an observation condition statement differs significantly from statements such as Q_1 and T_1. The difference seems to be this. Q_1 and T_1 are what we might call *first-order* observation condition statements, as they directly indicate that specific observation conditions obtain. Chisholm's Q, on the other hand, is what we might call, by contrast, a *second-order* observation statement. (Other instances of second-order observation condition statements might include such words as 'favourable' or 'unusual'). For in using the word 'normal' Q talks *about* first-order observation conditions. In particular, it refers to the whole class of first-order observation conditions and specifies that some are present and that others are absent. With respect to *C*, no first-order observation condition by itself seems to work. For as example 1 demonstrates, one first-order observation statement may counter-balance another. Example 2, however, strikes me as more plausible. But, if it is such examples of observation conditions upon which Chisholm's claims are based, then the phenomenalist hardly has cause to be encouraged. For Chisholm's Q should *not* be taken as a *pure* physical object statement.

This may be made clear in two ways. First, Q contains an appearance statement ('redness will be sensed') as the consequent of the conditional which is its second conjunct. Second, we should note that the phrase 'normal conditions' itself should *not* be interpreted to refer just to such purely physical phenomena as lighting conditions and anatomical descriptions of sense organs. Under such an interpretation example 2 as well may be shown to be incorrect. For $(P \& Q)$ is consistent with

(H) The observer is hallucinating yellow sense-data.

And $((P \& Q) \& H)$ does not entail R. To avoid this problem, the expression 'normal conditions' in Q needs to be understood to refer as well to the kinds of sensings that are likely to be produced. Q, therefore, should not be treated as a pure POS, but as a *mixed* statement, referring both to physical objects and to appearances.

Since the gap between POS and SDS seems only to be bridged by introducing such a mixed statement, Chisholm's example (so understood) is not at odds with his criticism of phenomenalism.

The problem that I see for Maloney's objections may be put as follows. Consider P′. If its conjunct Q′ is a first-order observation condition, such as T_1 or Q_1, then, as example 1 suggests, P′ will not entail S. If Q′ is a second-order observation condition, such as Chisholm's Q, then P′ will entail S. But since such a P′ is not a pure physical object statement, it does not pose a problem for Chisholm's argument. Similar remarks apply to P″ if constructed with some second-order observation conditions. Suppose P″ is restricted, however, to first-order observation conditions. Such a P″ faces the Russellian objection that it will not imply an S without at least the further premise that P″ contains *all* the relevant Q's. There then remains the worry as to whether such a class of Q's would include some statements about appearances. The preceding discussion suggests that it would. To dispel this presumption, the phenomenalist would require more than the schematic objections that Maloney provides.[1]

The University of Nebraska-Lincoln,
1010 Oldfather Hall, Lincoln,
Nebraska 68588-0321, U.S.A.

© ERIC RUSSERT KRAEMER 1984

---

[1] I am indebted to discussion with Charles Sayward.

ON MAKING OTHERS DO GOOD

By M. L. POKRIEFKA

DOŠEN has cogently argued (K. Došen, 'Is it good to make others do something good?', ANALYSIS 43.4, October 1983, pp. 199–201) that if we wish to maintain something like

(P) It is good to make others do good

it will be necessary to tack on certain restrictions to avoid perverse conclusions. (P) itself may need some reformulation. For, on one reading, (P) sanctions my forcing you at gunpoint to give blood to the Red Cross. This sort of unacceptable conclusion can be avoided by replacing 'make' in (P) by a less ambiguous term, perhaps 'persuade'. Such refinement will not, of course, obviate the need for Došen's further restrictions. Moreover, I will argue that some of Došen's restrictions, themselves, need restrictions if they are not to lead to perverse conclusions of their own.

To avoid the conclusion that it is good to leave others to do good when we could have done it ourselves, Došen suggests the following restriction on (P).

> (Q) Whenever we can do something good, we should not make others do it before we have done it ourselves.

But imagine a case where x knows of a friend y that y has done something wrong and that it would be good if y's wrongdoings were made public. Suppose x also knows that by talking with y, x can persuade y to make a public confession. According to (Q), x should tell the world about y's transgression without giving y the opportunity to do so.

There will be many other cases in which (Q) will lead to unacceptable conclusions; cases, for example, where one person is in a better position (in any number of respects) than another to do some good thing.

It is not at present clear to me how to amend (Q) so as to account for the cases I have considered. Došen has shown that something like (Q) is needed, and I believe I have shown that (Q) itself will not do. Though Došen's article clarified a number of other points, I am still uncertain just when to persuade others to do good when I could do it myself.

Michigan State University,
East Lansing, Michigan 48824, U.S.A.

© M. L. POKRIEFKA 1984

CORRECTION

In 'More on Empirical Significance' by M. L. Pokriefka, ANALYSIS 44.2, March 1984, on page 92, five lines up,

$$\Delta \cup \{T\} \not\models S'$$

should read

$$\Delta \cup \{T\} \models S'.$$

Philosophical Investigations

Edited by **D.Z. Phillips**

Philosophical Investigations publishes articles which discuss, extend or criticise philosophical approaches pioneered by J. L. Austin, Gilbert Ryle and Ludwig Wittgenstein.

Volume 7 No. 3 July 1984

Articles:

Scientific Realism and Ordinary Usage OSWALD HANFLING
Wittgenstein and a Causal View of Intentional Action DONALD GUSTAFSON
Neither Knowing Nor Not Knowing NEWTON GARVER

Reviews:

Robert Nozick, *Philosophical Explanations* ANTHONY PALMER
John V. Canfield, *Wittgenstein: Language and World*
Brian McGuinness (ed.) *Wittgenstein and his Times* T.S. CHAMPLIN
D.Z. Phillips, *Through a Darkening Glass* COLIN LYAS
R.G. Frey, *Killing and Suffering: Moral Vegetarianism and Applied Ethics* CORA DIAMOND
Peter Singer, *The Expanding Circle: Ethics and Sociobiology* R.A. DUFF
G. Stuart Watts, *The Revolution of Ideas* JOHN H. WHITTAKER

Published quarterly: January, April, July and October
Subscriptions Volume 7 (1984)
Institutions: *£25.00 (UK); £30.75 (Overseas); $58.50 (US); $71.70 (Canada)*
Individuals: *£11.95 (UK); £14.65 (Overseas); $28.00 (US); $34.30 (Canada)*
Single issues are available; prices on request.

Orders, with remittance, or requests for specimen copies and further details should be sent to:
Sue Dommett, Journals Department, Basil Blackwell,
108 Cowley Road, Oxford, OX4 1JF, England.

Basil Blackwell · Oxford · England

Philosophical Studies

An Internationl Journal for Philosophy in the Analytic Tradition

Editor-in-chief: JOHN POLLOCK ISSN 0031–8116
Associate Editor:KEITH LEHRER

Philosophical Studies, a journal devoted to work in analytic philosphy, publishes articles in all general areas of philosophical investigation, philosophy of science, epistemology, metaphysics, ethics, and so forth, including the history of philosophy. The main criterion of articles published is that they contain rigorous argumentation and precise analysis of key concepts. Articles in the journal are intelligible to philosophers whose expertise lies outside the subject matter of the article. It is intended that a diligent reader of the journal will be kept informed of the major problems and contributions of contemporary analytic philosophy.

Subscription Information
1984, Volumes 45–46 (6 issues)
Institutional rate: Dfl. 175,– / US $ 70.00 per volume including postage and handling
Private rate: Dfl. 60,– / US $ 24.00 per volume including postage and handling
Private subscriptions should be sent direct to the publishers.

D. Reidel Publishing Company
P.O. Box 17, 3300 AA Dordrecht, The Netherlands
190, Old Derby St., Hingham, MA 02043, U.S.A.

MIND

A QUARTERLY REVIEW
OF PHILOSOPHY
edited by D. W. Hamlyn
with the assistance of S. D. Guttenplan

Vol XCIII No 371 July 1984

Articles
- The Incorrigibility of the Cogito: J. HARRISON
- The Modal Ontological Argument: R. KANE
- Fragmentary Sense: P. CARRUTHERS
- An Alleged Difficulty concerning Moral Properties: J. C. KLAGGE
- Double Effect and Moral Character: N. RICHARDS
- A Meeting of Minds: A. PALMER

Discussions
- The Asymmetry of the By-relation: D. H. SANFORD
- Death and Bodily Transfiguration: W. R. CARTER
- Bodies, Corpses, and Chunks of Matter—a Reply to Carter: J. F. ROSENBERG
- Sceptical Confusions about Rule-Following: S. G. SHANKER
- On Bernard Harrison and Rigid Definite Descriptions: T. C. RYCKMAN
- Davies on Recent Theories of Metaphor: D. E. COOPER

Book Reviews **Books Received** **Notices**

Orders to: Journals Department,
Basil Blackwell,
108 Cowley Road,
Oxford OX4 1JF, England

Annual subscription:
Institutions: £14.50 or $32.50 (Canada $37.75)
Individuals: £10.00 or $22.50 (Canada $25.25)
Single issues: £4.00 or $9.00 (Canada $10.50)

Published for the MIND ASSOCIATION by BASIL BLACKWELL

METAPHILOSOPHY

Editor: Terrell Ward Bynum

Vol. 15 : 2 April 1984

A common occurrence: conflicting duties.	BERNARD H. BAUMRIN and PETER LUPU
Assuming the logically impossible.	ALICE AMBROSE and MORRIS LAZEROWITZ
Selfhood, modality, and philosophies of mind.	ARTHUR E. FALK
Theoretical models and the theory of sense-data.	THOMAS VINCI
Ethical theory, confidentiality, and professional ethics.	WILLIAM C. STARR
The two basic fallacies.	ANTHONY WESTON

1984 Subscription rates
Institutions: £28.00 (UK) £36.50 (Overseas) $65.00 (USA) $79.75 (Canada)
Individuals: £16.50 (UK) £20.20 (Overseas) $38.30 (USA) $47.00 (Canada)
Orders, with payment, to: Sue Dommett, Journals Department, Basil Blackwell, 108 Cowley Road, Oxford, OX4 1JF.

INQUIRY

Editor: Alastair Hannay

Vol. 27 (1984), No. 1

The Incoherence of the Cartesian *Cogito*	Leon Pompa
Hermeneutics, Transcendental Philosophy and Social Science	Mark B. Okrent
Conflicting Social Paradigms of Human Freedom and the Problem of Justification	Gerald Doppelt
Self-development and Self-management: A Response to Doppelt	Carol C. Gould
Human Freedom from a Democratic Socialist Point of View: A Reply to Doppelt	Mihailo Marković
On the Alleged Inconsistency, Moral Insensitivity and Fanaticism of Pacifism	Richard Routley
Feyerabend's Democratic Relativism	Steven Yates
Actions and Psychophysical Intimacy	Ted Honderich
Longino and Heidegger on Objectivity	Philip Lewin
Being and Saying (E. Tugendhat: *Traditional and Analytical Philosophy)*	John Llewelyn
Taking on Superior Beings (S.J. Brams: *Superior Beings: If they Exist, How would we Know?)*	Hannu Nurmi
Cheapening Science (K.D. Knorr-Cetina: *The Manufacture of Knowledge)*	Joseph Agassi

Subscription price:
Institutions: N.kr. 270,— $45.00 — Individuals N.kr. 140,— $23.00.

Published quarterly by Universitetsforlaget, P.O. Box 2959 Tøyen, Oslo 6, Norway — P.O. Box 258, Irvington-on-Hudson, New York 10533, USA.

RATIO edited by
MARTIN HOLLIS

Volume XXVI (1984) **June**

GEORGE SCHLESINGER	Possible Worlds and the Mystery of Existence
MARIE McGINN	Kripke on Wittgenstein's Sceptical Problem
N.M.L. NATHAN	Social Choice Theory versus The People
NICHOLAS UNWIN	Substance, Essence and Conceptualism
JONATHAN HARRISON	The Truth of Statements about Knowledge, and the Propriety of making them
MARTIN CURD	Showing and Telling: Can the difference between right and left be explained in words?
J.E. TILES	On a New Methodological Individualism
PHILIP PETTIT	In Defence of 'A New Methodological Individualism': Reply to J.E. Tiles

Subscriptions 1984

Institutions	£22.95 (UK)	£28.20 (overseas)	$59.95 (USA)	$74.00 (Canada)
Individuals	£15.50 (UK)	£19.00 (overseas)	$42.00 (USA)	$52.00 (Canada)

Orders to: Journals Department, Basil Blackwell,
108 Cowley Road, Oxford OX4 1JF, England.

The German edition of *Ratio* is published by FELIX MEINER VERLAG,
Richardstrasse 47, 2000 Hamburg 76, West Germany.

Published by BASIL BLACKWELL, Oxford

NOTES

The ANALYSIS Committee consists of: Chairman, P. T. Geach; Secretary, T. R. Baldwin; J. H. Benson, Margaret A. Boden, A. E. Fisher, Andrew Harrison, R. H. Holland, J. H. McDowell, D. H. Mellor, R. G. Swinburne, A. R. White, C. J. F. Williams, Peter Winch. This committee is responsible for appointing and advising the Editor and for the general policy of the paper.

SUBSCRIPTIONS. ANALYSIS is published by Basil Blackwell Publisher Ltd, 108 Cowley Road, Oxford OX4 1JF, England. Each volume comprises four numbers, each of at least 48 pages, appearing in January, March, June and October. Orders should be sent to the publisher. Subscription prices are as follows:

	U.K.	Overseas	U.S.A.	Canada
Institutions	£11.00	£13.50	$33.00	$39.50
Individuals	£7.00	£8.60	$21.00	$25.00

U.S. MAILING. Second class postage paid at New York. Postmaster: Send address corrections to: Analysis: c/o Expediters of the Printed Word Ltd, 515 Madison Avenue, New York, NY 10022 (U.S. Mailing Agent).

CONTRIBUTIONS. Articles submitted for publication should be addressed to Mr Christopher Kirwan, Exeter College, Oxford OX1 3DP.

Contributors are asked to note the following requirements. Articles should normally not exceed 3,000 words in length, including references and footnotes. They should be type written in double spacing on one side of the paper only; photographic copies are acceptable. Footnotes should be kept to a minimum and wherever possible avoided altogether. Inset quotations and formulae should be well spaced from the text preceding and following them. The spacing of symbols within formulae should be carefully attended to. Single quotation marks should be used, except for purposes of internal quotation and "scare" quotes. Discussion papers should be sent in as soon as possible after the appearance of the article to which they refer.

Typescripts of rejected articles will be returned only if self-addressed envelopes with correct postage are provided (in the case of overseas contributors, international reply coupons sufficient for airmail).

Proofs of accepted articles will be sent to authors for correction, together with information about free copies. Typescripts of accepted articles will be retained by the Editor on the assumption that authors have kept their own copies.

Apart from correction of printers' errors, only very minor changes in proofs will normally be accepted by the Editor, to avoid delay and expense in production.

1 NOV 1984

Typeset by Unicus Graphics Ltd, Horsham. Printed and bound in Great Britain by
Whitstable Litho Ltd, Whitstable

105
An13

ol. 44 No. 4 (New Series No. 204) October 19

ANALYSIS

Edited by
CHRISTOPHER KIRWAN

CONTENTS

Parsing 'if'-sentences V. H. DUDMAN
Fitch's factives J. J. MACINTOSH
'It would have happened already': on one argument for a first cause DAVID A. CONWAY
Sorabji and the dilemma of determinism PAUL RUSSELL
A dynamic model of temporal becoming STORRS MCCALL
Is motion change of location? PETER FORREST
Degrees of comparison MICHAEL CLARK
Wright *versus* Lewis on the transitivity of counterfactuals E. J. LOWE
Comment on Lowe CRISPIN WRIGHT
INUS conditions A. J. DALE
Goldstein on quotation D. K. BUCKNER
Explaining Donnellan's distinction — a reply D. E. OVER
Nozick and knowledge — a rejoinder B. J. GARRETT
A note on a response of Hornsby's E. J. LOWE
McGinn on Benacerraf GRAHAM SPINKS
Necessary agnosticism? ROBERT MCLAUGHLIN
Cause in perception: a note on Searle's *Intentionality* GREGORY MCCULLOCH
A paradoxical train of thought MARTIN HOLLIS
There is no set of all truths PATRICK GRIMM

BASIL BLACKWELL PUBLISHER · OXFORD

ISSN 0003-2638

PHILOSOPHY DOCUMENTATION CENTER

105
An 13,

rly Sis.

America [illegible] journal which publishes articles on any aspect of philosop[illegible] historical. The editorial policy is to publish work of high quality regardless of the school of thought from which it derives.

History of Philosophy Quarterly. A new journal which began publication in 1984. The journal will interest itself in papers that cultivate philosophical history in the spirit of *philosophia perennis.*

The Philosopher's Index: An International Index to Philosophical Periodicals and Books. An up-to-date quarterly index with abstracts from English language books and from more than 300 international philosophy and interdisciplinary journals. An annual cumulative edition is also published. In addition, two retrospective indexes have been published, i.e., the *Retrospective Index to U.S. Publications From 1940* and the *Retrospective Index to non-U.S. English Language Publications From 1940.*

Philosophy Research Archives: An International Scholarly Journal With a Microfiche Supplement. Sponsored by the APA, the CPA, and the PDC, is the journal of philosophy whose purpose is to provide a means of publication that is free of many of the disadvantages and constraints of journal publication. In addition, the *Archives* accepts articles, translations, monographs, and bibliographies.

Teaching Philosophy. Explores ideas about teaching and learning philosophy by publishing articles, discussions, and reports on new courses, methods, materials and their impact on philosophy. Textbook and audio-visual reviews are regular features.

Directory of American Philosophers. A handbook of philosophy covering the United States and Canada with information on faculties of philosophy, names, addresses, and specialities of philosophers, data on colleges and universities, philosophical societies, centers, institutes, journals (includes information on acceptance rates, refereeing, types of manuscripts, and other useful information), publishers, and data on fellowships and assistantships (types and dollar amounts) available in philosophy. Statistics on the profession. *The 1984-85 Edition has just been published.*

International Directory of Philosophy and Philosophers. A handbook of philosophy covering Europe, Central and South America, Asia, Africa, and Australia. This directory provides names, addresses, and specialities of philosophers as well as data on colleges, universities, philosophical societies, institutes, journals, and publishers of philosophical works.

Bibliographies of Famous Philosophers. Hardbound bibliographies have been published on Edmund Husserl, Jean-Paul Sartre, Alfred North Whitehead, Henri Bergson, Martin Heidegger, George Santayana, and Thomas Hobbes.

Dialog. A computerized search service which provides custom bibliographies with abstracts on any logical combination of philosophical topics.

Mailing Labels of approximately 9,700 philosophers in the United States and Canada are available as well as addresses by specialty, i.e., Ethics, Aesthetics, etc.

Journal Typesetting. Journals are economically typeset using computer-driven photocomposition equipment. Books, ads, and other documents can also be done.

Bibliography of Philosophy. The PDC, the U.S. national center selects books for review, and forwards the reviews to the editors in Paris.

Philosophy Documentation Center
Bowling Green State University
Bowling Green, Ohio 43403 USA

Collected Papers on Mathematics, Logic and Philosophy

GOTTLOB FREGE
Edited by Brian McGuinness

This book presents all Frege's mathematical writings, his reviews, fundamental articles on philosophy of language which prepared the way for his systematic works, his answers to criticism, his controversy over geometry, and all his final publications developing further his philosophy of language. Many are translated here for the first time.
416 pages, **£28.50** (0 631 12728 3)

Frege: Tradition and Influence

Edited by CRISPIN WRIGHT

The contributors to this book, a major assessment of Frege's thought and his legacy, include D. Bell, G. Currie, P. Carruthers, C. Diamond, M. Dummett, G. Gabriel, R. Hale, M. Luntley, G. McCulloch, J. McDowell, H. Noonan, J. Skorupski, J. Weiner and D. Wiggins.
270 pages, paperback **£7.95** (0 631 13831 5)

Probabilistic Metaphysics

PATRICK SUPPES

In this clear and authoritative book, Patrick Suppes argues against the classical assumptions of determinism. Denying the possibility of complete descriptions of events and laws he proposes instead a new theory of 'probabilistic' metaphysics.
256 pages, **£19.50** (0 631 13332 1)

The Roots of Knowledge

NATHAN STEMMER

This stimulating book advances a novel theory which explains the acquisition of our first subjective beliefs and their function as a basis for the later acquisition of objective knowledge. In addition, the theory makes it possible to solve the Hempel, Goodman and lottery paradoxes, and to clarify several other problems among them Hume's problem, the role of trial-and-error in scientific inquiry, and the problems of rationality, similarity, and universals. Finally, the theory is expanded to cover also the acquisition of our first beliefs about the semantic system of our language.
208 pages, **£17.50** (0 631 13328 8)

Basil Blackwell

Property and Political Theory

ALAN RYAN

Until a few years ago, arguments about the rights of property were central to political theory. Alan Ryan brings to life the cut and thrust of these arguments from Locke and Rousseau, through Kant and Hegel, to Mill and Marx, and gives a spirited and controversial answer to the question underlying the whole discussion: why did Mill and Marx get it wrong?
208 pages, **£15.00** (0 631 13691 6)

But is it Art?

B.R. TILGHMAN

The question of how we identify, describe and evaluate works of art has exercised philosophers through the ages. In this lucid, witty and thought-provoking book, Professor Tilghman explores the possibility or impossibility of constructing a proper philosophical theory of art.
208 pages, **£15.00** (0 631 13663 0)

The Legacy of Wittgenstein

ANTHONY KENNY

In these marvellously clear essays, Anthony Kenny provides an extensive commentary on Wittgenstein's work. Going beyond exegesis, he also gives a critique of other philosophers and theorists — Chomsky, Gregory, Anscombe, Teilhard de Chardin, for example — in the light of Wittgenstein's discoveries. A general introduction assesses the contemporary relevance of his thinking.
178 pages, **£17.50** (0 631 13705 X)

Wittgenstein on Meaning

COLIN McGINN

This book puts forward an original interpretation of the *Philosophical Investigations* and related works by Wittgenstein. It examines in detail what he actually wrote about meaning, understanding, rules and private language. Wittgenstein, it is argued, was opposed primarily to the 'inner process' conception of understanding, and his alternative view — that following a rule is a technique or custom — is elucidated.
160 pages, **£12.50** (0 631 13764 5)

Basil Blackwell

The Philosophical Quarterly

Vol.34 No.136 JULY 1984

Edited in the University of St. Andrews by an Editorial Board
Chairman of the Board of Editors: LESLIE STEVENSON
Executive Editor: GORDON GRAHAM *Reviews Editor*: ANTHONY ELLIS

Eternal Thoughts	*Peter Carruthers*
Fregean Thoughts	*Harold W. Noonan*
Frege's Platonism	*Bob Hale*
The Philosopher behind the Last Logicist	*Joan Weiner*
The Sense of a Name	*Michael Luntley*
De Re Senses	*John McDowell*
Frege, Sommers, Singular Reference	*Gregory McCulloch*
The Sense and Reference of Predicates: a Running Repair to Frege's Doctrine and a Plea for the Copula	*David Wiggins*
Frege's Metaphysical Argument	*Gregory Currie*
What does a Concept Script do?	*Cora Diamond*

Subscriptions for 1984

Annual Subscription	UK	Overseas	USA	Canada
Institutions	£22.50	£27.50	$50.00	$62.00
Individuals	£10.25	£12.50	$25.00	$30.00

Orders, remittances and enquiries should be addressed to: Journals Department,
Basil Blackwell Publisher Ltd., 108 Cowley Road, Oxford OX4 1JF, England.

The Philosophical Quarterly

Vol.34 No.137 October 1984

Edited in the University of St. Andrews by an Editorial Board
Chairman of the Board of Editors: LESLIE STEVENSON
Executive Editor: GORDON GRAHAM *Reviews Editor:* ANTHONY ELLIS

An Unsolved Puzzle about Knowledge	*Thomas Tymoczko*
'Ought implies can' and two kinds of morality	*John Kekes*
What does the private language argument prove?	*Oswald Hanfling*
Evolutionary Epistemology and Ontological Realism	*A.J. Clark*

DISCUSSIONS

The Failure of an *A Priori* Argument for Realism	*Anthony L. Brueckner*
Ackerman on Propositional Identity	*Rod Bertolet*
Humes Dialogue IX Defended	*Donald E. Stahl*
Lucas, Godel and Astaire: a rejoinder	*J.R. Lucas*

Subscriptions for 1984

Annual Subscription	UK	Overseas	USA	Canada
Institutions	£22.50	£27.50	$50.00	$62.00
Individuals	£10.25	£12.50	$25.00	$30.00

Orders, remittances and enquiries should be addressed to: Journals Department,
Basil Blackwell Publisher Ltd., 108 Cowley Road, Oxford OX4 1JF, England.

Philosophical Investigations

Edited by **D.Z. Phillips**

Philosophical Investigations publishes articles which discuss, extend or criticise philosophical approaches pioneered by J. L. Austin, Gilbert Ryle and Ludwig Wittgenstein.

Volume 7 No. 4 October 1984

Articles:

Knowing the Past CHRISTOPHER CHERRY
Mystery and Mumbo-Jumbo GORDON GRAHAM
Wittgenstein's Anthropology RICHARD BELL

Critical Notices

John R. Searle, *Intentionality* NORMAN MALCOLM
Ludwig Wittgenstein, *Last Writings on the Philosophy of Psychology* WILLIAM J. DEANGELIS

Review

Harald Delius, *Self-Awareness: A Semantical Inquiry* ILHAM DILHAM

Published quarterly: January, April, July and October
Subscriptions Volume 7 (1984)
Institutions: *£25.00 (UK); £30.75 (Overseas); $58.50 (US); $71.70 (Canada)*
Individuals: *£11.95 (UK); £14.65 (Overseas); $28.00 (US); $34.30 (Canada)*
Single issues are available; prices on request.

Orders, with remittance, or requests for specimen copies and further details should be sent to:
Sue Dommett, Journals Department, Basil Blackwell,
108 Cowley Road, Oxford, OX4 1JF, England.

Basil Blackwell · Oxford · England

Announcing the publication of

The Collected Papers
of
Bertrand Russell

Volume 7
Theory of Knowledge: The 1913 Manuscript
Edited by Elizabeth Ramsden Eames in collaboration with Kenneth Blackwell

Bertrand Russell's *Theory of Knowledge: the 1913 Manuscript* is the only book-length philosophical work of Russell which has not been published in complete form, and as such is an invaluable addition to our knowledge of Russell's thought. The period in which this book was written makes its publication particularly significant since it represents a transitional period in Russell's philosophy about which little has been known. The conjectures concerning the shift from his work of the pre-World War period of the *Philosophy of Logical Atomism* of 1919 can now be understood in terms of the difficulties Russell experiences with theory of knowledge under the onslaught of Wittgenstein's criticism. The unsolved problems concerning the status of proportions, the treatment of relations and the theory of judgement emerge starkly in this uncompleted but systematic work.

May 1984 314 pp 0 04 920073 9 Hardback £35.00

Reviews of The Collected Papers of Bertrand Russell:

'The editorial task has been discharged with exemplary thoroughness and efficiency. The entire series of volumes, when completed, will be indispensable to a thorough study of the intellectual development of one whose influence on the philosophy of his and our time has perhaps been greater than that of any other single individual.'

The Times Literary Supplement

'The editors of this publishing project are to be congratulated on what promises to be a magnificent achievement . . . a notable scholarly enterprise.'

New York Times Book Review

Full details from:

George Allen & Unwin (Publishers) Ltd
PO Box 18, Park Lane
Hemel Hempstead, Herts, HP2 4TE

George Allen & Unwin
London Boston Sydney

ANALYSIS 44.4 OCTOBER 1984

PARSING 'IF'-SENTENCES

By V. H. DUDMAN

I

COMPARE the natural interpretations of

(1) If Oswald did not shoot Kennedy, someone else did
(2) If Hitler invades England, Germany will win the war
(3) If Hitler invaded England, Germany would win the war
(4) If Hitler had invaded England, Germany would have won the war.

Contemporary wisdom contrasts the first two with the third and fourth, as 'indicative' and 'subjunctive' respectively. Grammatically, though, it is the first that differs from the other three. Under the natural interpretation of (1), the string of words in the scope of 'if' is parsed as a sentence: it is understood exactly as if encountered alone between full stops. Not so for (2), (3) or (4). For instance, 'Hitler invaded England' is not parsed as a sentence under the natural, future interpretation of (3), for if it was it would betoken something past. Nor, in particular, is 'Hitler invades England' parsed as a sentence under the natural interpretation of (2), as anyone will discover who essays a literal modus ponens. Rather, the entire 'if'-clause is parsed as merely a component of the principal predicate: Germany will-if-Hitler-invades-England win the war.

The point is that English generates her 'if'-sentences in more than one way. The natural interpretation of (1) is encoded into (1) by a different kind of encoding program altogether from the one that generates (2), (3) and (4) from *their* natural interpretations. Actually, there are just *three* basically different kinds of encoding routines at work generating 'if'-sentences, and my initial objective is to distinguish them.

II

The *first* kind of routine begins by separately generating two distinct sentences, and then puts the two together, prefixing one with 'if'. It is thus that the natural interpretation of (1) fetches up, in English, as (1).

A *second* kind of encoding program generates

(5) If Baby cries we beat him

from a generalization about occasions of Baby's crying. Construing 'Baby cries' as a sentence here is of course out of the question.

Like the second, the *third* kind of encoding program is one where the entire 'if'-clause is generated by an episode, merely, in the generation of the principal predicate. This third kind of program is easiest introduced as the one responsible for the following six combinations of formal tense and notional time in the 'if'-clauses it generates:

If *x* *V*-s ...	FUTURE; cf. (2)
If *x* *V*-ed ...	PRESENT or FUTURE; cf. (6), (3)
If *x* had *V*-en ...	PAST, PRESENT or FUTURE; cf. (4), (7), (8)

Figure A

To its ministrations we owe not only (2), (3) and (4) but also

(6) If Her Majesty was here now she would be revolted
(7) If Her Majesty had been here now she would have been revolted
(8) If the auditors had come tomorrow they would have found everything in order.

(Encoding routines of this third kind provide for a good deal of free variation. For instance, 'was' is everywhere replaceable ad libitum by 'were', and the 'if'-clauses of (3) and (4) are respectively replaceable, again ad libitum, by 'If Hitler was/were [Were Hitler] to invade England' and 'Had Hitler [Were Hitler to have] invaded England'.)

Now inevitably, when the encoding programs are different, the encoded messages are different: different programs consult different information in generating their output sentences. In the next three sections I try to describe, in turn, the kinds of messages encoded in my three ways. At the end, after a fleeting comparison in section VI of these three kinds of message, I seek to put 'contraposition' in its place.

III

When an 'if'-sentence is generated the *first* way, out of two component sentences plus 'if', each of those two sentences encodes its own separate message, its own 'proposition'. Indeed, besides whatever information it is that conditions the selection of 'if' itself, the total information encoded into the 'if'-sentence comprises just those two propositions, the 'dependent' proposition, let us say, and the 'principal' proposition.

The effect of the juxtaposition greatly varies:

(9) If the dowager had been angry at the abrupt leave of absence he took, she was mightily pleased at his speedy return [Thackeray, 1852]
(10) If you want them there are biscuits on the sideboard
(11) If the gate is locked Grannie must [may/might/should/could/can't/needn't] still be inside
(12) If Germany pushed Austria into war in 1914, it was because Germany believed a general European war to be inevitable
(13) The dog, if it was a dog, ran off
(14) If Toad looks funny in his bathing-suit then I, for one, want to see him
(15) If Sir Jasper murdered the librarian then Sir Jasper has already killed once
(16) If he had planned the staircase to exhibit the grace and distinction of his wife's carriage, it had failed of its purpose. The unhappy Mrs MacGregor soon developed the obesity which so often accompanies a life of enforced leisure and unremitting anxiety. She was incapable of descending a staircase without clutching at its railing [Thornton Wilder, 1967]
(17) If the latter was surprised and delighted, too, he concealed it like a diplomat [Wodehouse, 1931]
(18) If she can ride a bicycle she can ride a scooter
(19) If Tyrone beats his wife he is a cad.

Standardly, though – cf. (11) to (19) – the dependent proposition is understood as a *hypothesis*: by saying 'if', the speaker announces that he is treating the dependent proposition as true (or as a principle, etc.) regardless of whether it actually is. And then the overall message conveyed to the hearer varies with the *motive* she imputes to him for invoking this hypothesis.

Sometimes – cf. (15) to (19) – the speaker is understood to be adopting the hypothesis in order to use it as a premiss to infer the principal proposition from. In this special case we have what I take leave to call a *hypothetical* interpretation. Under a hypothetical interpretation, the sentence in the scope of 'if' specifies a hypothesis, and the principal clause announces a conclusion the speaker draws from that hypothesis.

Occasionally – cf. (15) – the conclusion is obtained very directly. But mostly it is drawn from the hypothesis taken together, as co-premiss, with some further *ground* or grounds. Sometimes – cf. (16) – the author of a hypothetical will weigh in with its ground while he is about it. Sometimes – cf. (17) – it is only as an oblique way of affirming its obvious ground that the hypothetical is volunteered at all. At other times – cf. (18) – the intended ground is less readily identifiable. Still, the speaker is supposed always to *have* any necessary grounds for his hypothetical, even if not called upon to disclose them. Indeed these grounds, whatever they are, are claims to which the speaker is *committed.* (Concerning an individual speaker's *commitments*, see Hamblin, C. L., *Fallacies* (London 1970) pp. 257, 263.) In sum, a speaker who asserts 'If *P*, *Q*' under a hypothetical interpretation intimates that he somehow concludes that *Q* from the hypothesis that *P*, perhaps because his commitments include some proposition *G* such that the inference with *G* and *P* as joint premisses and *Q* as conclusion is valid. The reader has doubtless long since observed that the natural interpretation of (1) is a hypothetical, with the proposition that someone shot Kennedy as its (most obvious) ground.

Other, nonhypothetical, messages encoded the first way include the natural interpretation of

(20) If a man beats his wife he is a cad.

The point – Frege's – about (20) is that it is generated by the same sort of program as gives us (19). It is just that in (20)'s case the program incorporates extra steps.

Meanwhile, (14) overthrows the grotesque conceit that 'then' in 'if'-sentences signals implication. In (13), which incidentally illustrates the point that the 'if'-clause can *interrupt* the principal clause, the hypothesis that the fugitive was a dog is invoked, not for the sake of argument, but rather to enable the speaker to get on with his narrative while leaving a detail unsettled. In (12), the hypothesis insures the speaker against being left with nothing to explain. And (11), whose principal clause purveys a subjective venture of the speaker, serves as a cautionary reminder that the burden of a component sentence need not always be a message of the true-or-false variety.

The natural interpretations of (9) and (10) provide exceptions to the general rule that a sentence in the scope of 'if' introduces a hypothesis. But then, 'if' is *misused* under the natural, 'Austinian', interpretation of (10): the pedant who protests that there are biscuits on the sideboard whether wanted or not is technically in the right. And as for (9), its natural interpretation involves a quite separate sense of 'if' – one had almost said a distinct *word* 'if'.[1]

IV

When an 'if'-sentence is generated the *second* way, the words in the scope of 'if' specify a condition, and the overall message is always a generalization of one kind or another about occasions of that condition's satisfaction. What determines the formal tense under these interpretations is the time, present, past or past past, of this generalization's validity:

---

[1] Actually, Thackeray's 'if' stands to *our* 'if' as 'while' in the sense of 'although' stands to temporal 'while', and as reason-giving 'as' and 'since' stand to temporal 'as' and 'since', and as reason-demanding 'when' stands to temporal 'when' (cf. 'How could you believe me when I said I loved you, *When* you know I've been a liar all my life?'). But that is the beginning of another story.

(21) If the alarm sounds I usually ignore [will often/can sometimes/may occasionally ignore] it

(22) In those days, if the alarm sounded I usually ignored [would often/could sometimes/might occasionally ignore] it

(23) Until that fateful night, if the alarm had sounded I had usually ignored it.

To clear heads – see e.g. Quine, W. V., *Word and Object* (M.I.T. 1960) pp. 170, 171 – the point was always clear, but it still needs insisting that these generalizations are never exponible Frege's way. Far from capturing the natural interpretation of (5),

$\forall t$ (if Baby cries at t, we beat him after t)

simply makes no sense, for 't' would have to range over elements of McTaggart's *A*-series, yet 'Baby cries ten minutes ago' is not an English sentence. Verbs are not general terms true of times. Not that Frege ever suggested they were.

In examples like

(24) If you tease a camel, it spits at you,

which is taken from Geach, P. T., *Reason and Argument* (Oxford 1976; see p. 89), we descry both kinds of generality at once: generality with respect to people and camels, exponible Frege's way, and generality with respect to occasions, which is not. Geach paraphrases (24) as 'As regards any man and any camel: if he teases it, then it spits at him', but fails to remark that 'he teases it' is still not parsed as a sentence here.

I should perhaps add that everything I have maintained in this section applies as much to e.g. 'when(ever)', 'wherever', 'as', 'before', 'as soon as', 'after', 'unless' and 'until' as to 'if'. What informational factor it is that conditions the choice of 'if' itself for the encodement of such generalizations is a question upon which I have ventured nothing. But then, my theme is the grammar of 'if'-sentences, not the meaning of 'if'.

V

When an 'if'-sentence is generated in the *third* way, there is, again, no sentence in the scope of 'if'. Analyses which proceed in terms of true and false 'antecedents' here are flat wrong: the 'if'-clause identifies nothing which might *be* true or false. Rather, the words in the scope of 'if' again serve to specify a condition:

If you run out of gin there's another bottle in the pantry

> If this solution turns green when I add the reagent, the deceased died of hyoscine poisoning

> Whether Empson will ever write any more poetry is not my business. If he does, it will be interesting to see whether the landslide in English literary taste has left us with a public capable of appreciating him [John Wain, 1957]

(25) If Grannie visits [had visited] the dentist again she will brush [would have brushed] her tooth beforehand
(26) If the enemy attacks [had attacked] we will be [would have been] ready for him
(27) If Chappell does not score [had not scored] the winning run, Gilmour will [would have]
(28) If the surgeon does [had] not operate[d] the patient will [would have] die[d]
(29) If the patient does [had] not die[d] the surgeon will [would have] operate[d]
(30) If the alarm sounds tomorrow I will [shall/can/may/must/should$_1$/ought to/needn't/daren't] ignore it
(31) If the alarm sounded tomorrow I would [should$_2$/could/might] ignore it
(32) If the alarm had sounded yesterday [tomorrow] I would [should$_2$/could/might] have ignored it.

Indeed, comparing (30) to (32) with (21) to (23), we discover the self-same condition specified in all six 'if'-clauses: that of the alarm's sounding. However, whereas the generalizations conveyed by (21), (22) and (23) are about actual historical occasions of this condition's satisfaction, the natural interpretation of (30) is evidently a claim about an *imagined* occasion of its satisfaction; and the same goes for (31) and (32). For, notoriously, (32) can be volunteered about the past by a speaker who acknowledges the condition to have remained, as a matter of historical fact, unsatisfied. And when it is the alarm's *future* sounding that is in question, how *can* the speaker be adverting to some actual occasion of it? In short, and quite generally, an 'if'-sentence generated the third way always invokes the *imagined* satisfaction of its subordinate condition: to formulate such a judgment, the speaker is obliged to conduct a *fantasy* in which the 'if'-condition is satisfied. (In a case like (29)'s, what is imagined is of course a condition's *non*satisfaction, but I shall be ignoring this complication.)

Since the same imagined satisfaction of the subordinate condition is likewise invoked by the natural interpretation of

> When the alarm sounds tomorrow I will [etc.] ignore it,

the special contribution of 'if' itself is again still to seek. And again I decline to be drawn: suffice it that, by saying 'if', the speaker advertises his fantasy as one *in which*, whatever else, the subordinate condition is satisfied.

Standardly – cf. (25) to (32) – the principal predicate (of which, remember, the 'if'-clause is merely a subsidiary component) incorporates a *secondary auxiliary*, either 'will', 'shall', 'can', 'may', 'must', 'should$_1$', 'ought', 'needn't', 'daren't', 'would', 'should$_2$', 'could' or 'might'. Furthermore, we discover in these principal predicates exactly six combinations of formal tense and notional time. I enumerate them here using 'will' and 'would' as paradigms, the former representing the first nine secondaries and the latter the other four:

will *V*	FUTURE; cf. (30)
would *V*	PRESENT or FUTURE; cf. (6), (31)
would have *V*-en	PAST, PRESENT or FUTURE; cf. (4), (7), (8).

Figure B

In these cases, that is to say, the principal clause's secondary auxiliary helps to propound some judgment or other concerning its subject's past, present or future *V*-ing, and my grammatical proposal is that what is marked by the formal tense is the temporal location, present, past or past past, of a point *v* of view from which this judgment is arrived at by thinking *futurewards*. This accounts, I need hardly observe, for the pattern disclosed by Figure B. It also accounts, I shall urge below, for the information about the order of events which is so much a part of messages encoded in this way.

At this point, however, I suddenly reduce the scope of our discussion, henceforth confining attention just to 'will' and 'would'. In line with this restriction, let us mean by a *conditional* interpretation of an 'if'-sentence, an interpretations where the principal predicate fits Figure B (to the extent even of literally containing either 'will' or 'would') and the 'if'-clause fits Figure A. In this special sense, the natural interpretations of (2), (3) and (4) are of course all conditionals.

Now, it has long been appreciated that conditionals can incorporate information concerning the order of events: see e.g. Downing, P. B., 'Subjunctive Conditionals, Time Order and Causation', *Proceedings of the Aristotelian Society* 59 (1959) pp. 125–40. *Usually* – cf. (28) to (32), and, of course, (2) to (4) – the 'if'-condition is very definitely understood as satisfied first. Indeed, it makes no sense to append 'beforehand' to e.g. (28), which therefore evidently cannot be generated from any interpretation whatever under which the patient's death antedates the surgeon's failure to operate. On the other hand, there are cases aplenty where the 'if'-condition enjoys no such priority: cf. (25) to (27). An adequate theory of conditionals has to account for all these facts.

My own attempt will be forseeable. Naturally the overall message conveyed to the hearer will vary with the motive she imputes to the speaker for certain juxtapositions (cf. section III); but in any case the verdict pronounced in the principal clause has to be arrived at by imaginatively projecting steadily *futureward* from v in a fantasy *in which* the 'if'-condition is satisfied. In line with this, the natural interpretations of (25), (26) and (27) will be seen to involve, respectively, anticipation of Grannie's own anticipation of the dentist's chair, anticipation of the enemy's assaulting an already fortified position, and anticipation of a winning run without a prior dismissal. But when it comes to (28), no comparable rationalization presents itself, whereupon the surgeon's imagined failure to operate is constrained to fall between v and the patient's projected death.

VI

This business of time order, now that it has come up, affords handy confirmation of the grammatical trichotomy urged in section II. We have observed that (28), generated the third way, imposes a definite order of events upon its interpretations. So too does (5), generated in the second; but this time the ingredient temporal information *pairs off* occasions of the two conditions' satisfaction: each time he cries, Baby is beaten *after that*. Messages encoded the first way, in sharp contradistinction, are under no comparable constraint. Thus alternative hypothetical interpretations, incorporating opposite orders of events, are equally accessible to

If the surgeon did not operate, the patient died,

one sustained by the ground that all patients subsequently died whom the surgeon failed, that day, to operate on, and the other sustained by the ground that the surgeon operated that day on every patient brought to him alive.

VII

There survives an antiquated notion of *contraposition*, conceived in innocent disregard even of the distinction between sentence and message, and a fortiori of the fact that it is the message that determines the sentence and not the other way round. Still, while the notion is irremediably obscure, the facts are plain enough; and the whole truth of the matter is this. Any ground which sustains a *hypothetical* interpretation of 'If P, Q' also sustains a hypothetical interpretation of 'If not-Q, not-P'. Hypotheticals are condensed arguments, meet to be upended.

Should the question arise of the relation between the natural interpretations of (28) and (29), the answer is that they have nothing to do with each other. Those who urge contraposition as a reason for thinking otherwise are invited to defend their prima facie absurd belief without invoking the magic word, and in particular to identify the two items – propositions? sentences? – they conceive themselves to be interchanging and negating. Unless there is a sentence in the scope of 'if', there can be no argument to upend.

VIII

And that, truly, is the pons asinorum: a conditional, for example, cannot be any sort of condensed argument, for where is the premiss? Indeed hypotheticals and conditionals are in my submission products of quite different styles of reasoning. A hypothetical is arrived at by arguing from propostion to proposition, a conditional by envisaging a developing sequence of events.[2] The speaker who ventures (1) is owning to some such argument as 'Someone shot Kennedy; Oswald didn't; therefore someone else did'; the speaker who ventures (2) or (3) or (4) is owning to a *fantasy* in which Hitler invades England and Germany goes on to win the war. His ability to draw conclusions he owes to his humanity, his ability to imagine consequences to his animality.

Macquarie University
North Ryde,
New South Wales,
Australia 2113.

©V. H. DUDMAN 1984

---

[2] I pursue this theme headlong in 'Conditional Interpretations of "If"-Sentences', *Australian Journal of Linguistics* vol. 4, no. 2 (forthcoming December 1984). Richmond H. Thomason has kindly drawn my attention to insurgent sentiments of similar tendency in his 'Conditionals, Time and Causal Independence' (mimeograph), where he calls for 'a conditional kinematics'. I seize this opportunity to salute Timothy Smiley's masterly exposition in his 'Hunter on Conditionals', *Proceedings of the Aristotelian Society* 84 (1983-4) pp. 169-77; see p. 175ff. Smiley resolves the terminological problem differently from me, calling the natural interpretation of (1) 'indicative' and those of (2), (3) and (4) all 'subjunctive'. On the significance of tense in 'if'-clauses, I have ventured exact submissions in 'Tense and Time in English Verb Clusters of the Primary Pattern', *Australian Journal of Linguistics* 3 (1983) pp. 25-44. Finally, I take great pleasure in acknowledging, with gratitude, the unfailing help of D. H. Mellor.

FITCH'S FACTIVES

By J. J. MACINTOSH

IN Fitch's 'A Logical Analysis of Some Value Concepts' (*Journal of Symbolic Logic*, 28.2, June 1963, pp. 135-42) we are presented with a very general form of argument concerning factives and semi-factives (hereafter, for simplicity of exposition, I shall use 'factive' to refer to both indifferently). Recent discussion (see, e.g., W. D. Hart, 'Access and Inference', *Proceedings of the Aristotelian Society*, supp. vol. LIII, 1979, pp. 153-65; J. L. Mackie, 'Truth and Knowability', Analysis 40.2, March 1980, pp. 90-2; T. Williamson, 'Intuitionism Disproved?', Analysis 42.4, October 1982, pp. 203-7) of the Fitch argument, however, has concentrated almost exclusively on one small aspect of it, the epistemological, and draws conclusions that the argument does not, and indeed cannot, really warrant. In this paper I want to look at the argument in its general form and see what lessons it in fact holds for us. In what follows I recast Fitch's argument for purposes of subsequent exposition.

I. THE ARGUMENT

Consider a monadic propositional operator D, which obeys the following three principles:

Distrib: $D(p \mathbin{\&} q) \rightarrow (Dp \mathbin{\&} Dq)$

Fact: $Dp \rightarrow p$

Poss: $p \rightarrow MDp$

In what follows I shall assume elementary propositional logic, and a fairly minimal modal logic (T would be more than enough) that contains, *inter alia*, a necessitation rule, and a definition that reveals 'M' and 'L' as duals.

As an exegetical note, Fitch introduces *Poss* only in connection with an ontological proof interpretation of the argument; Hart makes use of it, as does Williamson, for a knowledge interpretation. Fitch is explicit about the applicability of necessitation to *Distrib* and *Fact*: he is only interested in those D's for which *Distrib* and *Fact must* hold. He does not require this of *Poss*, though later versions of the argument (e.g., Williamson's) importantly and inexplicitly do. As we shall see, because of a crucial counterexample, *Poss* is an implausible addition to the first two in both its necessary and its merely assertoric guises, and attempts to evade the counterexample fail.

For the purposes of this argument, then, we are to think of D as doing duty for things like, 'a knows that ... ', 'a brought it about that ... ', etc. For D's which do obey the above three principles, Fitch offers us the following initially surprising result: that if p, then Dp. To show this we first show the impossibility of $D(p \mathbin{\&} {\sim}Dp)$:

(1)	$D(p \mathbin{\&} {\sim}Dp) \rightarrow (Dp \mathbin{\&} D\,{\sim}Dp)$	*Distrib*
(2)	$D\,{\sim}Dp \rightarrow {\sim}Dp$	*Fact*
(3)	${\sim}D(p \mathbin{\&} {\sim}Dp)$	1, 2, *PL*
(4)	$L\,{\sim}D(p \mathbin{\&} {\sim}Dp)$	3, *Necess*
(5)	${\sim}MD(p \mathbin{\&} {\sim}Dp)$	4, *Df*

This result is perfectly general and holds for any factive which distributes over conjunction. Given *Poss*, the desired result follows at once:

(6) $(p \,\&\, {\sim}Dp) \rightarrow MD(p \,\&\, {\sim}Dp)$ *Poss*
(7) $p \rightarrow Dp$ 5, 6, *PL*

If we can use necessitation on *Poss*, we can use it on (7) to get:

(8) $L(p \rightarrow Dp)$ 7, *Necess*

Given (7) and *Fact* we have $p \leftrightarrow Dp$, so any D to which these three principles apply will be in a sense redundant, though not necessarily uninteresting. 'Whatever is, is conserved by God, and whatever is conserved by God, is', is a principle which makes 'is conserved by God' a redundant operator for existents, though, historically, people who have accepted the principle have not found the operator uninteresting.

II. THE INTERPRETATIONS

How shall we interpret this argument? In particular, are there *any* factives which both distribute over conjunction and fit *Poss*? Here are three tempting but unsatisfactory possibilities:

(a) The deity operator: reading D as 'is conserved by God', or with a tenseless 'is', 'is (ultimately) brought about by God', the argument turns itself into an interesting form of the ontological argument, an argument in which, for arbitrary states of affairs, we move from the possibility of their having been brought about by God to the actuality of their having been brought about by God. Indeed, if we opt for the necessary version of *Poss*, this conditional will itself be necessary (line 8).

One of the many nice things about this argument is the simplicity of its deity, and of its starting point. We are considering God simply in his guise as creator, and need none of the usual trappings of ontological arguments: 'exists' does not need to be a predicate, we do not need to assume that God has 'maximal greatness', and so on. It is only necessary to grant that he might *possibly* have brought about the states of affairs we see around us, and this seems very little for even an atheist to grant.

(b) Fitch's argument is not, however, by any means purely a theological one. Let us look at it in its second analogy guise, as a proof of *ex nihilo, nihil fit*. Reading D as 'is caused to obtain' we move from 'if a state of affairs obtains it is logically possible that it was caused to obtain' to '(necessarily,) if a state of affairs obtains, it was caused to obtain'.

(c) As the previously mentioned discussions have in effect established, 'know' is, for a variety of reasons, a bad candidate for *Poss*, but what about *self*-knowledge? Suppose we restrict the range of our variable to present mental states, and read 'Dp' as 'a knows that p is the case' where p must be a present mental state of a. We must assume that these states are closed under both knowing and negation, which is implausible, but perhaps not wildly so, and then our conclusion – this time a Lockean one – follows: the mind is indeed self-revealing; whatever is in it, is known to be in it.

One of the virtues of the Fitch argument is that it shows us how much we give away in granting *Poss*, which is, in fact, an extremely strong principle. This has been somewhat masked in the literature to date, since *Poss* has tended to be smuggled into the discussion, which is why I have chosen to put it *into* the argument as an explicit theorem. Using an obvious modality bracketing convention to indicate the two alternatives, we have:

$$(L)(p \rightarrow MDp)$$

which, as we have seen, commits us to quite a lot, so there is no particular reason for the atheist or the quantum physicist or the anti-Lockean to allow it. Denying it, however, commits us, for the p in question, to the claim (depending on which form we're denying), either that it is *possible* that, for that p, the D in question necessarily doesn't hold, or, in the assertoric version, that, for that p, the D in question necessarily doesn't hold. Here, as always, the weaker the claim modally, the stronger its denial must be. (This is the point behind considering *Poss* in its non-necessitated version.) Concerning these denials, more later. First, though, let us look at the question that the rejection of knowledge, creation, and causality raises: is there *any* D for which it can plausibly be claimed that the three principles hold?

None *leap* to mind. If, for example, we try using a standard operator such as 'L' our result ($Dp \leftrightarrow p$) shows that adding *Poss* as an axiom to a system such as T, say, would (since *Fact* and *Distrib* hold for 'L' in T) make it collapse into PL and so, *a fortiori*, would make any higher system collapse. If we restrict our variables to past tense variables, then we could try a qualifier such as 'is now irrevocable', which is what both the mediaevals and ordinary language seem to mean by necessity. Here, however, the collapse of the operator is predictable simply in view of the restriction to past tense sentences. And, of course, once this restriction is lifted, *Poss* becomes inappropriate.

A final option is a version of Diodorean possibility. Diodorus tells us that the possible is that which either is or will be. And if we allow that whatever is necessary is possible, then the past is included too, in which case Diodorean possibility (in this extended sense) will fit all three principles. But what this seems to show, among other things, is why philosophers have not been terribly interested in Diodorean possibility, save as an historical curiosity: for such a Diodorean possibility would collapse uninterestingly into the assertoric mode in any normal use, so that *Fact* would become '$p \rightarrow p$', and *Poss* would become '$p \rightarrow Mp$'.

III. THE CONCLUSIONS

The argument, then, despite recent appeals to it, shows nothing directly about realism, or about verificationism, or about intuitionism, or even, it seems, about theology or quantum mechanics. But it points a worthwhile moral about possibility, namely, that we must beware of agreeing too readily that things are *possible*. When we think of possibility as epistemic possibility – 'it's possible for all we know that . . . ' – many things are possible, but we must not substitute for this possibility that strangely strong notion, logical possibility.

Ultimately Fitch's argument leaves us with the following: that for factives which distribute over conjunction, this disjunction holds: that, for any given p, (necessarily) $p \rightarrow Dp$, or else there is some p such that (possibly) $p \,\&\, L \sim Dp$. Which is simply to say, of course, either the earlier result holds, or *Poss* has exceptions. If, in any particular case, we dislike the first disjunct (which renders D redundant), we must hold the second: that (it is possible that) there is something for which the factive *necessarily* fails to hold.

In either form the second option (which includes '$L{\sim}Dp$') may look surprising. The argument has revealed to us that upholding goodness and truth in the form of atheism, quantum mechanics, and the non-immediacy of the mental seems to commit us to a stronger claim than we might otherwise have thought required. That, for example, (it is possible that) there is something which, *necessarily*, is uncaused, and similarly in the other cases, is something we must now hold, for otherwise we are forced to agree that *everything* is caused (created by God, etc.).

But this result is truly *un*surprising: for what it is that is necessarily not brought about by God (not caused, etc.), is simply the *conjoint* state of affairs that something both is the case and is not brought about by God (or that something both is, and is uncaused), and these conjoint states *cannot* be brought about by God (or caused) since these notions are, *ex hypothesi*, distributive factives. Fitch has proved (line (5) in the argument above) that, for distributive factives, $L{\sim}D(p \,\&\, {\sim}Dp)$, so the atheist (say) who holds that, *in fact*, there is some state of affairs p that God did not bring about will face with equanimity the notion that, for that *conjoint* state of affairs ($p \,\&\, {\sim}Dp$) it is necessarily true that God didn't bring *it* about, for that claim – $L{\sim}D(p \,\&\, {\sim}Dp)$ – is a theorem.

John Mackie suggested (*op. cit.*; I have changed Mackie's '*J*' to '*D*' for consistency) that we should look for *D*'s with 'innocent' interpretations, adding that 'by "innocent" I mean that for pretty well any proposition p it is possible that Dp and also possible that ${\sim}Dp$'. As Mackie of course noticed the crucial exception to his 'pretty well any proposition' is provided by '$p \,\&\, {\sim}Dp$': we don't want to allow that it's possible (in the theological interpretation, for example) that God might make it the case both that p, and that he didn't make it the case that p.

That this counterexample is not trivial (or, equivalently, that *Poss* is audaciously strong) may be seen by considering the consequence of attempting to evade the counterexample. We want to rule out '$p \& {\sim}Dp$' as an acceptable substitution instance in *Poss*. But it won't do to rule out *only* '$p \& {\sim}Dp$', since anything that leads to it will also be an exception. So we must try, as a weakened version of *Poss*, something like:

W(*eak*) *Poss*: $(p \rightarrow MDp) \vee (p \rightarrow \exists q(q \& {\sim}Dq))$

that is to say, *Poss* holds for all propositions except those which lead to the crucial exception. However, this weakening is in fact so extreme (and it's hard to see how anything less extreme would do), that the new principle is derivable from

$Lp \rightarrow p$

that is to say, *W Poss* has disappeared as a separate principle. And needless to say, our results no longer follow.

Here is an argument which shows the non-independence of *W Poss*:

(1) $L{\sim}Dp \rightarrow {\sim}Dp$ T
(2) $(p \& L{\sim}Dp) \rightarrow (p \& {\sim}Dp)$ 1, PL
(3) $(p \& L{\sim}Dp) \rightarrow \exists q(q \& {\sim}Dq)$ 2, EG
(4) $(p \& L{\sim}Dp) \rightarrow (p \rightarrow \exists q(q \& {\sim}Dq))$ 3, PL
(5) $(p \rightarrow MDp) \vee (p \rightarrow \exists q(q \& {\sim}Dq))$ 4, PL

We must conclude that this new form of the ontological argument fares no better than the old, and so, *mutatis mutandis*, for the other versions. Are there any interesting *D*'s that can plausibly be held to fulfil all three of the Fitch principles? Well, it is beginning to seem unlikely. It is not only that any *D* that did fulfil the principles would be vacuous in the sense that *p* and *Dp* would be truth functionally interchangeable, for – as we saw – such a *D* might be *interestingly* vacuous (as the deity operator would be if *Poss* were a possibility for it).

Rather, the point is that, for any plausible *D*'s, there will be *p*'s such that '$p \& {\sim}Dp$' will be true. But substituting a formula of this form into *Poss* yields a contradiction. And the attempt to weaken *Poss* enough to avoid the result destroys it. I conclude, then, that for any philosophically interesting *D*, the three Fitch principles may be treated as an inconsistent triad, and any arguments built upon them are bound to fail.[1]

University of Calgary, Calgary
Alberta, Canada T2N 1N4

©J. J. MacIntosh 1984

---

[1] I am grateful to B. F. Chellas for a number of helpful suggestions. Work on this paper was done while holding a University of Calgary Humanities Institute Fellowship.

'IT WOULD HAVE HAPPENED ALREADY': ON ONE ARGUMENT FOR A FIRST CAUSE

By DAVID A. CONWAY

I

IN real-life disputes it is not uncommon to hear the theist challenge the sceptic: 'If there is no God, then how did the world get here?' Dressed up a bit philosophically, the idea behind this challenge may become: 'Everything that exists in the natural world is caused to exist by something else. But since an infinite regress of such causes is impossible, there must be a first uncaused cause which is not a part of the natural world. And that uncaused cause is God, the creator of heaven and earth.'

However prominent its role in popular argument, it is difficult to find serious philosophical discussion of this sort of claim. Many advocates and critics of 'cosmological arguments' focus instead on some sort of non-temporal 'first cause'. Others argue in terms of some version of the principle of sufficient reason, rather than in terms of any principle of causality as such. When the argument for a genuinely *first* cause is mentioned at all, it is most often dismissed after only the most superficial discussion. The impression is left that the issues lack any particular historical importance and raise no philosophical difficulties worth genuine consideration.

William Craig has recently attempted to change both of these impressions. He presents a great deal of valuable information on the history of the temporal argument, and he mounts a detailed and vigorous defence of it.[1] Unfortunately, many of his arguments for the crucial premise that an infinite regress is impossible do little more than point out that infinite series have odd properties, and so it is 'hard to believe' that they really exist.[2] Since being 'hard to believe' for common sense can hardly be taken as the touchstone of the impossible, arguments of this form are far from conclusive. Craig does sketch one argument, however, that seems to be stronger. It is supposed to show that the hypothesis of an infinite regress leads not just to 'oddities' but to real contradictions. But Craig leaves the argument in such an undeveloped state that we can only guess how it might go at its most crucial point. Here I will suggest some ways in which it could go and argue that none of them results in a sound argument for the impossibility of an infinite regress.

---

[1] William L. Craig: 'Wallace Matson and the Crude Cosmological Argument', *Australasian Journal of Philosophy*, Vol. 57, No. 2 (1979), pp. 163-70; *The Kalām Cosmological Argument* (London: Macmillan, 1979); *The Existence of God and the Beginning of the Universe* (San Bernandino: *Here's Life*, 1979); *The Cosmological Argument from Plato to Leibnitz* (London: Macmillan, 1980).

[2] On p. 45 of *The Existence of God*... Craig rejects the possibility that there could be an infinite collection explicitly on the ground that that would have implications that he finds 'hard to believe'. Other of his arguments in effect have this form but not quite so explicitly.

II

In a bit of a variant on Sterne's story of Tristram Shandy writing his autobiography, Craig supposes that Shandy has been writing from eternity and is just now finishing the final page. But

> then we may ask, why did Tristram Shandy not finish his autobiography yesterday or the day before, since by then an infinite series of events had already elapsed? No matter how far along the series of past events one regresses, Tristram Shandy would have already completed his autobiography. Therefore, at no point in the infinite series of past events could he be finishing the book. [Also, it follows that] at no time in eternity will we find Tristram Shandy writing, which is absurd, since we supposed him to be writing from eternity.[3]

That is, supposing that he writes from eternity and finishes at some time, it follows that there is no time at which he finishes and that he has not written at all.

Put a bit more strictly, the argument is, I think, intended to go like this.

(1) Suppose Tristram Shandy writes the infinite number of pages of his autobiography and finishes today.
(2) Since an infinite number of pages has been written by today then an infinite number of pages had been written by yesterday, by the day before, by . . .
(3) Since an infinite number of pages had been written by yesterday, by the day before, by . . ., then Tristram Shandy will have finished (i.e., written the last page) by yesterday and by the day before and by . . .; that is, by any previous time.
(4) Since he has finished writing by any preceding time, then
    (a) there is no point at which he finishes
    and
    (b) there is no point at which he was writing.

Either (a) or (b) contradicts (1), and so it cannot be that Tristram Shandy has written an infinite number of pages and finished today. That is, it cannot be that an event occurs after having had an infinite number of predecessors. Thus, Craig concludes, there must have been a beginning.

This argument begins well enough. (1) is just the supposition that is to be shown absurd, and the conditional on which (2) rests ('*if* an infinite number of pages . . . ') is easily justified. Let *TS* be 'Tristram Shandy completes the last page of his autobiography', *−2* be 'Shandy completes the penultimate page of his autobiography',

---

[3] *The Kalām Cosmological Argument*, pp. 98–9. Cf. 'Wallace Matson . . . ', pp. 166–7.

−3 be 'he completes its third from the last page', etc. The series of events, of page-writings, would then be represented by

. . . *−5 −4 −3 −2 TS.*

If we assume that he writes at the methodical pace of one page per day, then the day before yesterday he will have written page −3, and so at the end of that day altogether he will have written the series of pages,

. . . −7 −6 −5 −4 −3,

which is itself clearly an infinite number of pages. Indeed even though it lacks two members that the series culminating in TS has, the two series can easily be seen to have the same number of members since for each member of either series there is a corresponding member of the other,

. . . −5 −4 −3 −2 TS
. . . −7 −6 −5 −4 −3.

Thus, the first is 'no more infinite' than the second.

Since similar reasoning will apply if we remove from the first series not two but two thousand or two billion or $2{,}000^{2{,}000}$ members (pages written, events), then no matter how far back in time we may trace the series, at that point it will already have an infinite number of members. And that is what (2) tells us.

III

It is when we arrive at (3) that things become extremely puzzling. For (3) is hardly self-evident, no argument is offered for it, and none presents itself as an obvious and defensible candidate. What is needed is a rationale for the conditional that lies behind (3), i.e., a rationale for

(E) If an infinite number of pages had been written by yesterday, by the day before, by . . ., then Tristram Shandy will have finished by yesterday and by the day before and by . . .

One naive but perhaps appealing rationale could come from thinking of the infinite as 'the all-encompassing'. 'There can be', one might have heard it said, 'nothing outside the infinite'. Or, 'the infinite cannot be only partial; it must include everything'. If so, any infinity that has transpired by yesterday, or by the day before that, or by . . . must include everything. It must include, then, Tristram Shandy finishing his autobiography.

It may be that there is some (metaphysical? Spinozistic? Hegelian?) sense of 'the infinite' in which it is 'the all-encompassing', but if so, it would be quite an incredible confusion to think it relevant here. For the present concern is with the infinite in the sense in which a series may have an infinite number of members and that clearly has nothing to do with being 'all-encompassing'. The whole number series is infinite even though it encompasses no werewolves, harpsicords, or theologians. The series

4 5 6 7 8 ...

is infinite even though it does not include even all the positive integers. And a series of Tristram Shandy's page-writings may be infinite without encompassing the writing of the first page of *Ulysses* or the last page of Tristram Shandy's autobiography.

But perhaps the hidden rationale for (E) is not that an infinite series must 'include everything' and so must include some event TS. Rather, the idea could be that given that a series is infinite and *does* include TS, then it would not be infinite without it. Just as a ten-member series does not 'include everything' but would not be a ten-member series if we took away one of the members that it does have, an infinite series need not include everything but would not be an infinite series if it lacked one of the members that it has. And so, since by hypothesis (in (1)) Tristram Shandy's infinite series of page-writings as of today includes writing the last page, then if he has written an infinite number of pages by yesterday, by the day before ..., then he must have written the last page by any of those times. Without TS the series just would not be infinite.

The apparent plausibility that this rationale has rests on thinking of infinite series as if they were finite ones. Of course if you take some members away from a finite series you have left a series with fewer members, but, as we have already seen in regard to (2), the number of members of an infinite series is unaffected by the removal of any finite batch of members.[4] And so it cannot be necessary for the series being infinite that Tristram Shandy's page-writings include his writing the last page. The number of members in the series of pages written will have been just the same without the last page, or the page before that, or ... having been written.

IV

In another place Craig says this:

> Suppose we meet a man who claims to have been counting backward from eternity and now he is finishing: ... −5, −4, −3, −2, −1, 0. Now this is impossible, for we may ask, why didn't he finish counting yesterday or the day before or even the year before? By then an infinity of time had already elapsed, so that he should have finished.[5]

---

[4] Or by the removal of some infinite batches, but this fact has no direct bearing here.

[5] *The Existence of God ...*, p. 48.

From here, the argument proceeds as above. That is, the unfortunate consequences in (4) are drawn in the same way.

This argument looks different at first glance. Before the claim was that the occurrence of an infinite number of *events* would require the occurrence of a specific event (TS). Now the idea is that an infinite period of *time* requires that some event occurs. But this surely is a difference without any substance. If The Counter must have reached the last of the infinite number of numbers, it is not because an infinite period of time has passed *per se*. (The Counter could, for all we know, have remained silent throughout such a time.) It can only be because an infinite number of numbers are taken to have been enunciated when an infinite period of time has passed; i.e., an infinite number of events are taken to have occurred. Thus, no real advance has been made over the formulation in (1)-(4).

V

Taking a clue from a footnote in Craig,[6] we might try to support (E) and (3) by appealing to a sort of 'possibility principle'.

(P) All possibilities are actualized at some time during an infinite period of time.

Or, less scholastically,

(P) In an infinite period of time, anything that can happen does happen at some time or other.

There is some plausibility to this principle, if we think of it in terms of a certain kind of model. If 20 fair dice are being rolled, one time after another, then the more times they are to be rolled the more likely it is that a certain total, say, 98, will be rolled. Given an infinite period of time and so an infinite number of rolls, it seems quite inevitable that 98 and any other possible number will be rolled. Generally, there is a good deal of intuitive appeal to the claim that, given long enough, each thing that can happen will happen, an infinite period of time must be 'long enough', and so, at any time at which an infinite period of time has passed, any particular event will already have occurred.

Again the argument appears to be about what will happen in an infinite period of time, but the justification for (P) in terms of the number of rolls of the dice makes it obvious that what matters is that an infinite number of *events* has taken place. What (P) really comes to is:

---

[6] Here I must confess that Craig's footnote cites approvingly an earlier work of mine in which I use (P) to generate an 'it would have happened already' claim. My aim, though, was to give a sort of *reductio* of what I take Aquinas to be arguing in the 'Third Way', and so some things (e.g., (P)) were assumed there for that purpose. At least some of the appropriate questions about those assumptions are given above in Section V. (See David A. Conway, 'Possibility and Infinite Time: A Logical Paradox in St. Thomas' Third Way', *International Philosophical Quarterly*, Vol. XIV, No. 2 (1974), pp. 201-8.)

> (P′) If an infinite number of events had taken place by any particular time, then any possible event will have taken place by that time.[7]

And it is this that is now to serve as a justification for (E) and so for (3) in the basic argument of (1)-(4).

But (P′), even if true, would seem to be quite unsuited for this role. For, most plausibly construed, it will tell us only that during an infinite period of time there occur such events as *the number 98 is rolled, someone called 'Horace' is born,* and *a philosopher devises the ontological argument.* That is, certain 'sorts' of events occur, certain event 'types' which can happen as often as one might be pleased to imagine. And there will now be no problem in the supposition that one of these happens as a result of an infinite series of events having transpired. For the problem is that an event happens at some time but before any particular time, and so at no time. But events that can happen over and over can happen both at some time and before any particular time. There is nothing even a little bit mysterious in the claim that it is snowing now and that it had snowed before today, before yesterday, before ..., before any particular past time.

We might try to fix this up by asserting that (P′) refers to 'logically nonrepeatable' events such as *the birth of the first man ever called 'Horace'.* But if we do this, (P′) loses its initial plausibility, for there is just not much intuitive appeal to the claim that a specifically *first* (or a second or a third ...) anything must occur in a beginningless series of events.

Or perhaps we could try to fix up these problems with (P′) in yet another way, one that would bring out that it is not just a curiosity dreamed up here for current intellectual amusement. In the 'Third Way', the argument of St. Thomas most generally beloved by contemporary Thomists, Aquinas may well be arguing that in the absence of a necessary being it is possible that it should come about that nothing exists, and so, in an infinite time with an infinite number of events, this will have come about.[8] For Thomas, the absurd implication is that this will have happened before now and so nothing exists now. The absurdity is not that the event *it comes to be that nothing exists* happens before now *and* before any other time and so at no time. Also, what Thomas takes to be the offending premise and so the one that he proceeds to reject is that there is no necessary being, rather than that there has been an infinite regress of events. But whatever the differences, what is important here is, first, that Aquinas apparently relies on (P′) in claiming that it *will* come about that nothing exists; and, second, this argument gives us a good candidate for a possible event that is not *logically* nonrepeatable (it is not, for instance, specifically a *first* occurrence of ...) but that nonetheless can happen only once. For when it has happened that *it comes about that nothing exists*, there must forever thereafter be nothing in existence, and so it can never again *come about that* nothing exists. Since the event is not *logically* nonrepeatable, (P′) should apply to it. And given both that (P′) applies to it and also the hypothesis of an infinite number of events, it must have occurred at some time and yet before any particular time. But since it is nonrepeatable in some sense, the impossibility is now generated: this event occurs at some time and yet at no time at all.

---

[7] (P′) bears a superficial resemblance to the rejected claim that the infinite is 'the all-encompassing'. But (P′) is not saying that the infinite as such requires that everything be included. It is saying that given the occurrence of an infinite number of events together with certain assumptions about how things take place in the world, then every possible event will be included in an infinite number of events.

[8] Copleston and many others so interpret the 'Third Way'. (F. C. Copleston, *Aquinas* (Baltimore: Penguin, 1961), p. 120.)

Unfortunately, understanding the basic argument in this way involves increasing its complexity more than its real plausibility. One problem with it is that it is difficult to see how the event *it comes about that nothing exists* can really be both possible and nonrepeatable. If what is in question is logical possibility, then it is possible that it comes to be that nothing exists, but it is also possible that something comes to exist again ('from nothing') after which it could again come about that nothing exists. But if it is objected that it is not *really* possible for something to come to exist from nothing, then it must be replied that it is not *really* possible for everything to cease to exist. (Some sort of principle of conservation of matter has been generally accepted since Aristotle.) It is, in short, logically possible that this event happens as many times as one might be pleased to imagine, but it is not really possible for it to happen even once. In neither case is it both possible and nonrepeatable.

The most bedrock difficulty with this version of the argument is that (P′) is, whatever its intuitive plausibility, quite a wrongheaded claim. For that intuitive plausibility rests on our somehow being beguiled into thinking of the world as a sort of giant dice game in which events happen 'at random'. And just making such a model explicit is enough to refute it, for one need not be a 'determinist' in any but the most commonsensical and commonplace sense to reject the view that we live in a world of randomly occurring events. As an explanation of the fact that a six was rolled with a single fair die we would accept that that was because it was a possible outcome and there were a lot of rolls. But we would never accept an explanation of why an earthquake occurred that said only that the earthquake was a possible event and a lot of events had taken place. For we take it that we live in a world of causal connectedness, a world in which events happen because there are sufficient antecedent conditions for their happening. The 'random model' is not one that we are willing to accept as a reflection of how things are. The rationale for (P′) depends on such a model, and so there is no reason whatsoever to accept (P′).

VI

We have looked rather carefully here at the claim that it is not possible that an event happens at some time after an infinite number of events have passed, because, at any point in an infinite series of events, 'it would have happened already'. We have not found any rationale for it, and it is now extremely difficult even dimly to imagine how such a thing might go. On the other hand, it is hardly less difficult to imagine how one could give a demonstration of the principled impossibility of there being any such rationale. Grand 'Kantian' manoeuvres are seldom successful. Specific arguments need specific consideration.

University of Missouri-St Louis,
8001 Natural Bridge Road,
St. Louis, Missouri 63121, U.S.A.

©DAVID A. CONWAY 1984

SORABJI AND THE DILEMMA OF DETERMINISM

By PAUL RUSSELL

IN *Necessity, Cause and Blame* (London: Duckworth, 1980) Richard Sorabji attempts to develop a notion of moral responsibility which does not get caught on either horn of a well known dilemma. One horn is the argument that if an action was caused then it must have been necessary and therefore could not be one for which the agent is responsible. The other horn is the argument that if the action was not caused then it is inexplicable and random and therefore not something which the agent can be responsible for. Sorabji denies that what is caused is always necessitated. Causes are primarily explanatory rather than necessitating. This established, Sorabji hopes to show that action open to moral scrutiny may be caused without being necessitated and the dilemma collapses. I will argue that this strategy runs into serious difficulties.

I

Sorabji argues that desires and beliefs are causes of action, and clearly he means to suggest that they are *efficient* causes of action (cf. p. 40). These desires and beliefs, it is implied, have weights which may or may not be 'equally balanced' when they oppose one another (I use the term 'weight' here to correspond with Sorabji's metaphor of 'scales', cf. pp. 30–1). Sorabji offers two examples in support of his claim that human action may be caused without being necessitated. In Chapter Fourteen he considers the case of a child who has equal but opposing desires. On the one hand the child wants to take another child's toy which appeals to him but on the other hand he equally wants to please and obey his parents. As it happens the child takes the toy but '*in the very same circumstances* the child could have acted in the other way' (p. 232n – my emphasis). We can, however, give an explanation and cite a cause of the child's action; namely, 'that set of incentives which favoured taking the toy'. Thus, while there is 'no answer available for someone who ... wants to be told why the child acted on the one set [sc. of incentives] rather than the other' there is nevertheless an explanation for why the action occurred. In this way, it is argued, the problem of randomness may be overcome.

In Chapter Two Sorabji offers an example 'in which the case for and against a decision does not seem equally balanced'. An agent may have 'overwhelming' reason for attending a lecture and 'nine times out of ten in such circumstances a man may act accordingly'. However, the agent also has some reason for not attending the lecture and therefore may sometimes not attend. (If 'scales' are balanced at a ratio of 9:1 they *always* go to the weightier side – not at a ratio of 9:1. How, given the analogy, does Sorabji account for the exceptional case? If the 'scales' do not always come down on the weightier side how do we establish the relative weight of rival sets of incentives? This sort of difficulty suggests that the metaphor of 'scales' is perhaps not a very happy one.) In other words, Sorabji suggests that under identical circumstances C an agent may usually do X (i.e. attend a lecture) for given reasons, say R_1, but he may occasionally do Y for other conflicting reasons, say R_2. Hence under the said circumstances neither X or Y was necessitated, though doing X was more probable because R_1 was weightier than R_2. Nevertheless, whether X or Y occurs they will be explicable; X in terms of R_1 and Y in terms of R_2. However, a serious problem remains and one which curiously does not seem to trouble Sorabji very much.

Sorabji admits that what cannot be explained is why the agent does X (acts on R_1) rather than Y (acting on R_2) or vice versa. 'But then ought we to expect', he continues, 'that there will be an explanation available corresponding to every contrast that we care to choose?' (p. 31). Perhaps we need not be able to answer the question which asks why an agent does one given action rather than *any* other action. But there is another distinct question which does seem more pressing: given that the agent has a set of *actual* rival desires and beliefs (R_1, R_2, . . . R_n) why, in any given case, does the agent act on some and not others? Let us call this the *comparative question*. In effect the comparative question asks if the agent can determine which set of incentives become his will. (The notion of will I use here is that of an *effective* set of incentives – one that moves the agent to undertake an action.) Sorabji seems to think not, and in this he certainly parts company with traditional libertarianism. What matters, it seems, is that when we do act there is some set of incentives which explain the action but do not necessitate it.

II

It seems clear, then, that the crucial issue is whether or not Sorabji is right in claiming that we need not be concerned about being unable to explain why an agent acts on one set of incentives rather than others which are also actually his. That is to say, can we ignore the comparative question and dismiss it as irrelevant to the issue of moral responsibility?

Let us assume that the difference between X and Y is morally significant. Action X, let us suppose, is some morally repugnant action (e.g. pushing an old lady in front of a bus) explicable in terms of a given set of incentives, R_1 (e.g. a dislike of old people, and a belief that this lady is one such person). Action Y, by contrast, is some morally commendable action (e.g. helping the old lady across the road) explicable in terms of another distinct set of incentives, R_2 (e.g. a love of one's fellow human beings, a pleasure in helping people and a recognition that help was required). Under conditions C the agent has overwhelming reason to do Y and, let us assume, 99 times in 100 would act on R_2. Despite this he does X and R_1 thereby becomes his will. Sorabji's argument suggests that because X was not necessary, is explicable in terms of R_1, and the agent could categorically have done Y he is (ceteris paribus) responsible. It does not matter, so it seems, that the agent had other much stronger and wholly different beliefs and desires which he did not act on. Nor does it matter, apparently, that he was unable to determine whether he acted on R_1 or R_2 – in other words, determine whether R_1 or R_2 became his will.

It may be argued that if the agent did in fact act on R_1 rather than R_2 then he must have opted for R_1 in preference to R_2 – thereby determining X in preference to Y. That is to say, when we act on R_1 rather than R_2 there must have been some 'new rival temptation or force' which 'tips the scales' in favour of R_1. This assumption of the determinist is, Sorabji claims, simply a 'declaration of faith' (pp. 30–1). We require no new rival temptation or force to distinguish cases where we do X from those where we do Y. In identical circumstances (i.e. C) either R_1 or R_2 could be efficacious. There is no further influence of the agent which accounts for the occurrence of one action rather than the other.

Consider the following contrast. Suppose that in given conditions C we have two sets of incentives, R_1 and R_2. In one case both R_1 and R_2 are causally determined and R_1 in turn necessitates action X. In the second case R_1* and R_2* may cause but cannot necessitate actions X* and Y* respectively – either alternative being quite possible. Given that X* occurs it will have been caused but unlike the first case (i.e. X) it will not have been necessitated. Sorabji would deny that the first case is an instance of morally free or responsible action because although the action was caused by the agent it was 'necessary all along'. How, then, would he regard the second case? As an action that was caused by the agent but not necessitated he would seem to regard it as one for which the agent is responsible. There is, however, no greater scope for agency in this case than in the first. The agent cannot determine whether he acts on R_1 or R_2 and so he no more determines which set of incentives becomes his will in this case than in the first case. That is to say, the opening up of possible actions in the second case does not increase the agent's influence over what he does. In both cases the agent fails to act on R_2 (or R_2*). In the first case necessarily, in the second case because of chance. Accordingly, the lack of necessity in the second case does not contribute to the agent's moral freedom.

What is disturbing about Sorabji's account of the agent's conduct in these circumstances is that it suggests that what the agent does *within the range presented by C* (i.e. R_1, R_2, ... R_n) is simply a matter of chance. The *nature and character* of the actions which the agent produces in these circumstances will be as random and subject to chance as the *'rate and route'* of radioactive emissions (Sorabji, it should be noted, uses this analogy between radioactive emissions and human conduct; cf. pp. 28–30). While it is certainly true that such actions are produced by the agent's will, it is equally clear that the agent can no more determine which set of incentives he will act on (i.e. which course of action he will pursue) than 'the lump of radioactive material' can determine the rate and route of its emissions. Thus, clearly in these circumstances it is not up to the agent how he *wills*.

It may be suggested that Sorabji's position can be defended against this line of criticism if we simply assume that 'a man is responsible (excusing conditions aside) for what he does by will, regardless of the origin of his will'. However, for Sorabji's purposes this argument proves too much; it would serve equally well as a defence of the compatibilist position (a position which Sorabji explicitly rejects). That is, if we adopt this defence for Sorabji's position then an agent could be justifiably held responsible for an action which was *necessitated*, as long as it was the agent's will that brought that action about.

The above considerations establish that acting on R_1 does not of itself adequately eliminate the problem of randomness as Sorabji thinks. He has limited the range of randomness (i.e. by excluding those actions with no causal connection with C and by establishing a causal connection between the action in question and C) but within C the randomness persists. In other words, within the range presented by C ($R_1, R_2, \ldots R_n$) it is simply a game of roulette as to what action will follow. That I do X rather than Y, and therefore that I am condemned rather than praised, becomes a matter of chance (whether the 'odds' are 1:100, 1:10, or 1:2 does not matter).

III

There is no denying that at first glance Sorabji's attempt to resolve the dilemma of determinism would seem to satisfy the demands of both libertarians and compatibilists. On the one hand his account gives libertarianism the categorical freedom (i.e. the possibility of doing otherwise in identical circumstances) which it seeks while, on the other hand, it meets the compatibilist demand that action be caused by the (antecedent) willings of the agent and not mere random happenings. However, despite this initial plausibility, I believe that both the traditional libertarian and the orthodox compatibilist would argue that Sorabji's theory suffers from serious defects. Their dissatisfaction, I suggest, is indicative of the fact that, for different reasons, neither party would accept that the comparative question is irrelevant to the difficulties which the dilemma of determinism presents us with.

From the traditional libertarian point of view Sorabji's thesis suffers from the same fatal flaw as the compatibilist's position. Namely, on Sorabji's account the agent cannot determine which of the pre-established alternative sets of incentives he acts on. While libertarians, notoriously, find it difficult to give much coherent content to the notion of a non-phenomenal 'self', or the notion of how the will is 'determined' by this 'self', it is nevertheless quite clear why they believe that we need to appeal to these (obscure) notions. Moral responsibility, they argue, requires *moral autonomy*; and moral autonomy requires *self-determination*.

The traditional libertarian cannot accept that the comparative question is irrelevant to the issue of moral responsibility because it is clearly relevant to the issue of moral autonomy. For the traditional libertarian an agent is responsible in so far as *he* (i.e. a non-phenomenal 'self') determines his *antecedently* undetermined will. Thus the explanation for why an agent in circumstances C does X rather than Y is that he (i.e. the 'self') *chose* to act on that set of incentives. To be able to answer the comparative question is to establish that such an agent had moral autonomy and was capable of self-determination. Hence, to dismiss the comparative question as irrelevant to the issue of moral responsibility is to dismiss the issue of moral autonomy and self-determination as irrelevant to the issue of moral responsibility.

From the orthodox compatibilist perspective Sorabji's thesis suffers from a serious defect common to all libertarian theories: it allows too much scope for the play of chance. Given Sorabji's account of the explanation of human action agents may frequently and inexplicably act contrary to their dominant inclination. This claim raises a number of difficulties from a compatibilist point of view. In particular, in these circumstances how could we accurately infer anyone's character or disposition on the basis of his actions? If agents frequently and inexplicably act out of character and contrary to their stronger inclinations we would find ourselves unable to anticipate their future actions reliably. In general, it may be argued that human motivation is more consistent and more coherent than Sorabji's model suggests. Indeed, it may be psychologically impossible to survive in a world in which people are randomly acting on rival sets of incentives. (Imagine how difficult it would be to live in a world in which everyone acts as if they had multiple personalities.) In this way, it seems clear that Sorabji's thesis suffers from the same sort of difficulties which compatibilists have found in other, traditional, libertarian theories.

For the compatibilist the only way in which the unacceptable play of chance can be eliminated is if the strongest set of incentives is the effective set of incentives and vice versa. As long as this is the case we can answer the comparative question. That is, the explanation for why an agent in circumstances C does X rather than Y is that the effective set of incentives (i.e. R_1) was stronger than the alternative set of incentives available to him (i.e. R_2).

In short, it is not evident that Sorabji's thesis succeeds in avoiding all the major difficulties associated with the dilemma of determinism. Indeed it may be argued that in some important respects his theory gets impaled on *both* horns of the dilemma at once. At the heart of these difficulties lies the issue of whether or not the comparative question can be dismissed as irrelevant to the issue of moral responsibility. Sorabji *assumes* that it is irrelevant. However, as we have noted, to assume this is simply to beg some of the key issues that are in dispute. At the very least, therefore, we require independent arguments in support of this assumption.

In my view the dilemma of determinism is not likely to be resolved by Sorabji's suggestion that what is caused need not be necessitated. However, the fact that Sorabji's rather ingenious attempt to resolve this dilemma runs aground on the same rocks as its predecessors is itself of considerable interest. The regular failure of this sort of project suggests that philosophers approach the issue of moral responsibility with the wrong set of questions in mind. If we want to avoid running aground on these rocks we must chart a new course – we must ask different questions. Most philosophers, Sorabji included, approach the issue of moral responsibility by way of trying to provide some adequate account of how a responsible agent 'could have acted otherwise'. Their problem then becomes: how are those *actions* for which we are justifiably held responsible *caused*? The intractable nature of the dilemma of determinism suggests that there is something fundamentally wrong with this entire approach.[1]

Sidney Sussex College,
Cambridge

© PAUL RUSSELL 1984

---

[1] I am grateful to G. A. Cohen, B. A. O Williams, S. Makin, the Editor, and especially to Richard Sorabji for helpful comments on this paper.

A DYNAMIC MODEL OF TEMPORAL BECOMING

By STORRS MCCALL

PROFESSOR J. J. C. Smart has kindly commented on a theory of time contained in a paper of mine.[1] The issue is whether there is such a thing as the flow or passage of time, or whether time flow is a subjective illusion. Does time 'of itself, and from its own nature', in Newton's words, flow 'equably, without relation to anything external'? Or is the correct view of the universe that of a four-dimensional continuum, spread out in three dimensions of space and one of time, in which every event from the deepest past to the most distant future finds its place?

In such a world nothing 'flows'; were there no conscious observers there would not even be the illusion of time flow. The question of whether such flow is indeed a subjective illusion, or whether there is a real analogue to time flow, is the question Smart and I are interested in.

To be clear about what is at stake, consider the four-dimensional Minkowskian picture of the universe. The thesis that time is *anisotropic* is the thesis that there are physical processes which are directionally biassed, i.e. which occur always or for the most part in one temporal direction rather than in the other. For example, if you light a match outdoors on a clear, starry night a spherical shell or wave-front of photons begins to expand outwards in space.

---

[1] Storrs McCall, 'Objective time flow', *Philosophy of Science* 43 (1976), pp. 337-62; J. J. C. Smart, 'Time and becoming', in *Time and Cause*, ed. Peter van Inwagen, Dordrecht 1980, pp. 3-15.

Expanding photon-shells are frequently encountered, but what is never or almost never seen is a contracting shell of inward-moving photons, a wave-front that in theory could ignite a bare matchstick held out in the dark.[2]

Irreversible processes confer an asymmetric texture or 'grain' to the universe along the time axis, but they do not, as Grünbaum points out, single out one of the two possible temporal directions as *the* direction of time. The match experiment does not designate the direction of increasing values of the time-coordinate as *the* direction of time, since an equivalent description of the phenomenon would be that contracting spherical wave-fronts almost always occur in the direction of decreasing time-coordinates. Hence, though time is anisotropic neither direction need be preferred.

Anisotropy by itself, therefore, does not serve to indicate the direction of time. Although the Minkowski world exhibits a grain along the time axis, it does not come labelled 'This side up'. Such labelling requires the additional assumption that one of the two time directions is preferred. This assumption is not equivalent to the assumption that there is temporal becoming, though it is implied by it, and so I am not able to say what Smart's views on it would be. Let us leave the question for the moment and turn to the more controversial matter of whether time passes, with the consequent division of events into 'past', 'present', and 'future', and the consequent labelling of the two temporal ends of the universe (assuming it has ends) with 'start' and 'finish'.

For Grünbaum, and for Smart, temporal passage is mind-dependent in the sense that without the successive entry into consciousness of features of the world there would be no such thing as 'now-ness' or the present, consequently no constantly changing demarcation of past and future. A world without conscious observers would be a world without time flow, a world in which people 'grow' ten years older only in the sense in which a highway 'runs' ten miles north of Canberra.[3] But is the mind-dependence theory correct? Is time flow an illusion? What I tried to do in my 1976 paper was to produce an objective correlate of time flow: a feature of the universe which had nothing to do with whether conscious observers existed or not, a feature which could model temporal passage. The theory, which makes essential use of the notion of physical possibility, is as follows.

The Minkowski world consists of a single four-dimensional manifold containing every event that has been, is, or will be. Large though this collection of events is, it omits future events that, though physically possible, will not occur. Relative to a given set of initial conditions, two or more different outcomes are frequently possible. It is possible for Tom to win the 6/49 lottery, and possible for him to lose. It is also physically possible for the temperature of the interior of the sun to go up a millionth of a degree, and possible for it to go down. These possibilities are not epistemic, reflecting only what we know or do not know about lotteries or the sun, but are 'real' or 'objective' physical possibilities. Combining them in all different ways yields four possible futures: (i) Tom wins and the temperature goes up, (ii) Tom wins and it goes down, (iii) Tom loses and it goes up, (iv) Tom loses and it goes down. At most one of these possibilities will be realized, which is what I had in mind in saying that there are physically possible future events that will not occur. Such events can be arranged in branched four-dimensional manifolds which themselves branch continuously towards the future; each path through the tree represents a possible history of the world.

---

[2] The example is Adolf Grünbaum's, from whom I have learned almost all I know about time. For the distinction between anisotropy and temporal becoming see Grünbaum's *Modern Science and Zeno's Paradoxes*, Wesleyan U.P. 1967, p. 9 ff.

[3] Smart, p. 5. Cf. also Paul Davies: 'We cannot escape the conclusion that the qualities of time we regard as the most vital – the divisions into past, present and future, and the forward movement of each division – are purely subjective. It is our own existence that endows time with life and motion. In a world devoid of conscious observers, the river of time would cease flowing.' (*Other Worlds*, New York 1980, p. 46).

In contrast to the single four-dimensional manifold of the Minkowski world, which has the topological structure of a straight line, the branched arrangement of possible world histories has the structure of a tree, with a single trunk representing the past, the first branch point representing the present, and a multitude of branches representing the various possible futures. There is nothing fanciful or imaginary about this tree. Every detail of the mode of branching, and the precise nature of every event on every branch, is as much an empirical matter, dictated by what physical laws permit or fail to permit relative to given initial conditions, as the structure of Saturn's rings. Of the many possible futures relative to 31 January 1985, one and only one will become actual, but no physical mark presently indicates *which*.

Let us call the whole tree-like branched structure of four-dimensional manifolds 'the universe'. Then the dynamic feature of the universe which I suggest as the physical analogue of time flow is this: the successive shedding of 'unused' branches of the tree as the first branch point (the present) moves upwards. Making allowances for special relativity,[4] trees corresponding to later times are proper sub-trees of trees corresponding to earlier times, i.e. 'older' universes have a different topology from 'younger' ones.

Smart feels that this proliferation of universe-trees, one for each instant of time, places too high a strain upon our willingness to tolerate ontological plurality. We would have, as he points out, continuum-many trees each with continuum-many branches.[5] But the matter can be looked at in another light. Instead of a different tree for each time, we can think of the different trees as 'snapshots' of the universe at different times. The snapshots look different, showing that time is passing and the universe is growing older. But it's the same universe throughout, just as different snapshots of little Sophie reveal the same person at different stages in her life. Sophie is a dynamic entity, as witnessed by the diverse appearance of her photos at different times. In the same way the universe is a dynamic entity, as is revealed by its different tree-structure at different times. This is why the theory is described as a 'dynamic' theory of time flow.

---

[4] The world-instant designated as the 'present', and hence the branching above it, is determined not only by a time but also by a frame of reference.

[5] Smart, p. 7. In actual fact, the number of branches on each tree is even greater than Smart's figure. See 'The cardinality of possible futures', *The Journal of Symbolic Logic* 35 (1970), p. 363.

There are, of course, differences between the universe and Sophie. Sophie is a three-dimensional being with temporal extension, or, which comes to the same thing, a small branched four-dimensional object within a large branched universe. The universe, unlike Sophie, is not embedded within anything larger. But in other respects Sophie and the universe are similar. As the universe changes, Sophie changes. Or perhaps we should say, as Sophie changes, the universe changes, since many of Sophie's changes are *her* doing, not the universe's. Both Sophie and the universe change in two different senses of the word 'change'. First, just as the three-dimensional Sophie has long hair at one stage and short hair at another, so the three-dimensional universe has lots of red giants at one stage and lots of white dwarfs at another. Second, just as the four-dimensional universe has more branches at an earlier stage and fewer branches at a later stage, so four-dimensional Sophie has more branches earlier and fewer branches later. Like Sophie, the universe is a dynamic entity. Its dynamic character serves as the physical analogue of time flow.

Perhaps enough has been said to indicate how the dynamic theory differs from the mind-dependence theory. There still remains a big question of course, namely which theory is correct? It's not my intention to argue the point here. I like the dynamic theory, and feel in my bones that it does more justice to our ordinary beliefs about passage of time than the mind-dependence theory. But I am content at this stage merely to present the two theories as internally consistent alternatives and let the reader choose. I'm not sure how Smart feels about this: – perhaps he will argue that the dynamic theory is *not* an internally consistent alternative. We shall see. In conclusion I want to deal with an important but subsidiary point he raises in his paper.

On page 8 Smart questions my assertion that nothing now indicates which of the many future branches of the universe-tree is going to become part of the trunk (i.e. which is *the* future) as opposed to which are going to be lopped off. For Smart, my position is too verificationist. He would maintain, presumably, that though nothing may now *indicate* which branch is going to be selected, still there is one and only one branch of which it is *true* that it will be selected. Of all the events on that branch it is true that they will occur, whereas it is true of at least one event on every other branch that it will not occur. In this way Smart might argue that we are able to reduce the universe to a single branch after all, via the notion of 'what will be', and the truth of future-tense sentences.

When I wrote my 1976 paper I was dead set against the idea that one of the future branches could be singled out in any way, even by the notion of truth, and went to considerable lengths to construct a 'temporal' theory of truth which allowed many sentences about the future to be neither true nor false. Now I believe I could have saved myself the trouble. It no longer seems to me objectionable or wrong-headed to allow that there is a specific person now walking the streets of Canada who will be our next-but-one Prime Minister, and of whom it is true to say that he or she will be. To say this in no way implies, for example, that it is impossible for Tom Smith or anyone else to become the next-but-one PM. No doubt this is possible for any of about 10 million people. Nothing *physical* now marks out any of them as the lucky one, and the truth of the statement that *X* will be the one is entirely dependent upon (supervenient upon) what will happen ten or so years from now. The notion of truth, so to speak, bakes no bread, it simply floats on top of whatever events occur or will occur and in no way constrains or affects the possibility of any of them occurring. This 'supervenient' character of truth is no threat to the branched universe, and in particular does not succeed in reducing the number of branches to one, any more than it succeeds in reducing the number of potential prime-ministerial candidates to one. What is true depends upon what will exist and what will occur, but what will exist and will occur does not depend upon what is true.

To sum up, contrary to what I said in 1976 I no longer see any objection to the class of all true propositions about the future 'picking out' one of the future branches as *the* future. Such picking out results from a one-way dependence relation from future branch to true proposition, rather in the style in which a footprint 'picks out' a murderer. What is picked out by true future-tense propositions is what the universe *will be* when all its days are run. What it *is* remains the dynamic universe-tree.

McGill University,
Montreal, PQ, Canada H3A 1G5

© Storrs McCall 1984

IS MOTION CHANGE OF LOCATION?

By PETER FORREST

THE current orthodoxy about motion[1] is that:

(1) X moves (absolutely or relatively) just in case X's (absolute or relative, respectively) location changes.[2]

In this note I argue, by means of counter-example, that (1) needs to be modified, *at least*. I say 'at least' because the complexity of the required modification suggests that we might seek a radically different analysis.[3]

I begin with the counter-example of a disc, cone, sphere, cylinder or any other object with an axis of symmetry. Suppose it is homogeneous and that it spins about its axis. Then its location does not change. But it spins. Therefore it moves. So change of location is not necessary for motion.

The obvious response to this counter-example is to point out that many of the *parts* of the spinning object change their location. This suggests my first modified analysis:

(2) X moves just in case there are parts of X which change their location, and which make up all or most of X.

There is, however, a counter-example to both (1) and (2). It concerns the Achilles-weed and the tortoise. Achilles-weed is prostrate and grows along the ground at the amazing rate of 10 cm per hour. An exceeding slow tortoise munches one end of the Achilles-weed at the same rate as it grows at the other end. So the tortoise appears to chase the Achilles-weed round the garden. But, strictly speaking, the Achilles-weed does not move at all, it grows and is eaten. Yet its location changes, and it is made up of parts whose location changes (the left and right-hand halves of the Achilles-weed). Hence being made up of parts whose location changes is not sufficient for motion.

Notice that although the Achilles-weed *can* be divided into parts which change their location, it can also be divided into parts which do not. In order to provide an analysis that is proof against the Achilles-weed example I first make a definition. I say that Y *almost homœomerously changes location* if, however, we divide Y into parts, those which change location make up all, or most, of Y.

---

[1] See, for example, M. Jammer, 'Motion', *Encyclopedia of Philosophy*, *Volume Five*, ed. Paul Edwards (Collier-Macmillan 1967), p. 396.

[2] We might insert the qualification 'continuously'. I am not prepared to judge whether this is required. If it is, it can be inserted in all the analyses I shall frame.

[3] I have in mind a *fluxus formae* concept of motion involving a quality of absolute instantaneous velocity, or a dyadic relation of relative instantaneous velocity.

My next modified analysis is proof against both counter-examples discussed.

(3) X moves just in case X is made up of parts which almost homœomerously change location.

However, solid objects at a temperature above absolute zero are made up of *vibrating* molecules. And a vibrating molecule almost homoeomerously changes location. Yet not all solid objects are moving. We require a further modification, so as to obtain:

(4) X moves just in case X is made up of parts, of size comparable to X, which almost homœomerously change location.

Australian National University,
Post Office Box 4, Canberra ACT 2600,
Australia

© PETER FORREST 1984

DEGREES OF COMPARISON

By MICHAEL CLARK

CAN comparative sentences be analysed without quantifying over degrees? In his comment on Adam Morton's paper 'Comparatives and Degrees' (ANALYSIS 44.1, January 1984) Christopher Williams (*ibid*) says he would like to find a simpler analysis of sentences like 'Sally has grown fatter' than the form offered by Morton, viz.

> $\exists d \exists e \exists t$($a$ is F to degree d now and a is F to degree e at time t and d exceeds e in $\mathscr{L}$(F) and t is earlier than now).

$\mathscr{L}$(F) is the characteristic partial ordering of F, and the whole analysis is presented as a more rigorous and explicit rendering of

> $\exists g \exists h$(Sally's girth is now g and Sally's girth was h and g is greater than h).

Williams is understandably reluctant to accept that growth can be as complicated as Morton's account makes it. What leads Morton to analyses which involve quantification over degrees is a desire to offer a general schema which serves for sentences like 'Roger is fatter than we thought possible' and 'Dennis drank more than the legal limit' as well as simpler cases like the sentence about Sally; and it is difficult to see how anything less complicated than Morton's proposal would meet this requirement. But perhaps a more satisfactory account can be found if we try setting aside the requirement.

After all, if you sought a common schema for sentences like 'Blair was Orwell' and 'Blair was an author', you would probably construe the latter as having the form '*b* was identical with some author', which, though not incorrect, is less simple and illuminating than the form '*Fb*'.

Following Williams, let us paraphrase the sentence about Sally's growth as:

Sally is fatter now than she was.

We can bring out its formal relations with sentences like 'Sally is fatter than Tony (is)' and 'Sally is fatter than Tony was' by treating it as expressing a four-place relation between a person, a time, a person and a time:

x is fatter at t_1 than y at t_2.

For reasons elucidated by Geach ('Some Problems about Time', in *Logic Matters*, Blackwell, 1972) I think we should avoid treating the sentence as expressing a two-place relation between time-slices of Sally (Sally-now and Sally-then). The four-place relation is logically prior to girth and its degrees, since the latter are unintelligible without it. We could not have a concept of girth without knowing what it was for one creature to be fatter than another, whereas we could talk about Sally growing fatter even if we had not abstracted the concept of girth or its degrees. Given the comparative notion, girth can be defined in terms of the relation, x has the same girth at t_1 as y at t_2, which in turn can be defined as: x is at least as fat at t_1 as y at t_2, and y is at least as fat at t_2 as x at t_1.

I have used the expression 'at least as fat as' rather than 'no fatter than' in the definition of girth. This is to allow for incommensurable cases, where it is indeterminate whether one girth is greater than another. (Even if girth cannot be indeterminate in that way, other qualities can, as Morton illustrates with his artificial example of independent ecclesiastical and temporal hierarchies on p. 19.) If we take 'x is at least as fat at t_1 as y at t_2' as relatively primitive, then 'x is fatter at t_1 than y at t_2' can easily be defined in terms of it (x at least as fat at t_1 as y at t_2, but y not at least as fat at t_2 as x at t_1).

Quantifying over degrees can be avoided in the analysis of another of Morton's examples, 'John ran faster than the Olympic record': we simply paraphrase it as 'John ran faster than anyone else had run in the same Olympic event'. It is when we come to the sentences 'Roger is fatter than we thought possible' and 'Dennis drank more than the legal limit' that his style of analysis seems difficult to avoid. But then these sentences do make fairly obvious reference to degrees, namely how fat we thought it possible (for Roger?) to be and how much the law allows motorists to drink. Yet, although girth is definable in terms of the comparative notion, the definition cannot be used to eliminate quantification over degrees in these more complex cases. And as soon as comparatives fall partially within the scope of psychological, modal and analogous expressions, we seem to have reached a level of abstraction at which degrees are indispensable. It is perhaps even more obvious that we have reached that level with sentences like 'Sally has grown twice as fat as she used to be' and 'Cyril is many times fatter than Sally'.

The complexity of the analysis required should not surprise us in these later cases, but we are not forced to recognize the same complexity wherever comparison is involved. There is an analogy here with numerical sentences. 'There is exactly one apple in his basket' can be expressed in the symbolism of pure first-order predicate logic supplemented with identity: $\exists x(Ax \,\&\, \forall y(Ay \rightarrow x = y))$. But to translate sentences like 'There are twice as many apples in her basket as in his', 'There are three times as many &c.' into distinct forms we need to quantify over numbers or sets or properties. We can, of course, provide a more complicated analysis of the first sentence to parallel these analyses, construing it as 'The number of apples in his basket is one', but the surface structure of the English sentences does not tempt us, as it does in the case of Morton's examples, to discard the simpler analysis as inadequate.

University of Nottingham,
University Park, Nottingham NG7 2RD

© MICHAEL CLARK 1984

WRIGHT *VERSUS* LEWIS ON THE TRANSITIVITY OF COUNTERFACTUALS

By E. J. LOWE

IT is part of contemporary orthodoxy in the logic of conditionals that for subjunctive conditionals, including counterfactuals, inference by transitivity is not valid: see especially David Lewis, *Counterfactuals*, Blackwell 1973, pp. 32 f. This doctrine has recently been challenged by Crispin Wright, in his 'Keeping track of Nozick', ANALYSIS 43.3, June 1983, pp. 134–40. However, while I am sympathetic with Wright's objective, I am not happy with his manner of pursuing it, as I hope to make clear in what follows.

Wright remarks that if he were to assert

(1) If there had been snow in the valley yesterday, I would have gone skiing

we are not at liberty, putting this alongside the uncontentious

(2) If an avalanche had then been taking place, there would have been snow in the valley yesterday

to conclude

(3) If an avalanche had been taking place yesterday, [I] would have gone skiing

'For', he says 'it is at least likely that . . . I did not intend to be avowing my willingness to have gone skiing in suicidal circumstances' (op. cit., p. 137). However, Wright does not (as Lewis might) take this example as illustrative of the non-transitivity of subjunctive conditionals, properly understood. For his position seems to be that if he were to assert *both* (1) *and* (2) 'in the same breath' (as he puts it) we *should* take him to be prepared to assert (3) also, though not if he were to assert (1) alone (despite the fact that (2) is an uncontentious item of common knowledge). The lesson he wishes to draw is that the Lewisian semantics for counterfactuals are in error for failing to recognize that the 'acceptability' (Wright's surrogate for truth) of an asserted counterfactual cannot be assessed independently of what other counterfactuals are (as he puts it) 'in play in the same context'. And in consequence he holds that 'subjunctive conditionals, like all conditionals, sustain transitivity in context' (op. cit., p. 140).

There are two main reasons for my unhappiness with Wright's proposal. First, I don't think he does enough to persuade a supporter of Lewis to concede that if he *were* to assert both (1) and (2) 'in the same breath' he *would* thereby be committing himself to (3). All he says on this score is this: ' . . . it is clear that in any normal conversational context, you would be bound to wonder — wouldn't you? — whether I had meant to hint that I had been depressed, or some such thing' (op. cit., p. 138). However, Lewis seems entitled to suggest that if a 'hint' *is* all that is being made, his theory could explain this more adequately than Wright's precisely by *denying* that an entailment is involved: he could say that the hint is effected *pragmatically* simply by the juxtaposition in speech of two counterfactuals whose assertion 'in the same breath' would not otherwise serve any normal conversational purpose, since one of them expresses an item of common knowledge.

Secondly, and more fundamentally, Wright's proposal seems to me to have dire implications for logic. For Wright, as we have seen, wants to relativize the *validity of inferences* involving counterfactuals to conversational contexts. It will, on this view, no longer be proper to ask whether certain premises entail a certain conclusion without consideration of which of those premises (if any) are to be regarded as being asserted 'in the same breath' as which others. I need hardly say how much this would complicate logic! It seems reasonable to object that Wright is simply obliterating the distinction between semantics and pragmatics.

But, in any case, an adherent of transitivity need not go to Wright's lengths to cope with the avalanche example. Certainly we may agree with Wright (and Lewis) that someone asserting (1) on its own would not normally be supposed to be committing himself to the truth of (3). But why not say that this is only because such a person does not *literally* mean to assert (1), but rather something which he could more literally (if more pedantically) express by

(4) If the snow conditions in the valley yesterday had been suitable for skiing, I would have gone skiing

That such a person really meant (4) rather than (1) could readily be elicited by the question 'Would you then have gone skiing if there had been an avalanche, or if there had been only a thin scattering of snow? – to which the answer would naturally be 'No, of course not: I only meant that I would have gone skiing if the snow conditions had been suitable'.

However, here I anticipate some such objection as the following on behalf of the Lewis camp. A person might want to assert (4) and *also* to assert

(5) If there had been snow in the valley yesterday, the snow conditions would have been suitable for skiing

But given transitivity, (4) and (5) seem to entail (1), while by transitivity again (1) and (2) entail (3). So it seems that such a person *is* committed to the truth of (3) after all, if transitivity is valid (unless, of course, one takes Wright's view). Worse still (apparently), by transitivity (2) and (5) entail, it seems,

(6) If an avalanche had been taking place yesterday, the snow conditions in the valley would have been suitable for skiing

But in fact it is not too difficult for an adherent of transitivity to turn the tables on the Lewisian here. For he may point out that a devotee of Lewis might himself be well advised *not* to regard (5) as a standard counterfactual interpretable in the usual Lewisian way. To see why we need to recall that according to Lewis, a counterfactual is true if and only if amongst the worlds in which the antecedent is true, the consequent is true in those worlds 'closest' to the actual world. But if (5) is interpreted along these lines it is difficult to explain why anyone should ever consider that sentence (non-vacuously) true. For on a very natural understanding of 'closeness', given that there was in fact *no* snow in the valley yesterday, the worlds 'closest' to the actual world amongst those in which there *was* snow would seem inevitably to be worlds in which there was only a thin scattering of snow, quite *un*suitable for skiing! A world in which there was only a little snow in the valley surely must be – other things being as nearly equal as possible – 'closer' to a world in which there was none than is a world in which there was a substantial amount. Of course it might well be that, owing to the valley's peculiar climate, a world in which there was only a thin scattering of snow would be, in some sense, highly *improbable* (perhaps in this valley it rarely snows but it pours) –

but it need be none the less 'close' for all that: unless, of course, 'closeness' is precisely to be *understood* in probabilistic terms – in which case one might as well go straight for a probabilistic semantics in the first place (something which Lewis himself does not favour, op. cit., p. 71).

Much the same point may be made another way, as follows. If a devotee of Lewis insists on regarding (5) as a standard counterfactual, he cannot explain why transitivity should *fail* in the inference from

(7) If there had been a thin scattering of snow in the valley yesterday, there would have been snow in the valley yesterday

and (5) to

(8) If there had been a thin scattering of snow in the valley yesterday, the snow conditions would have been suitable for skiing

For, of course, the antecedent of (7) is *not* 'more far-fetched' than that of (5), as the Lewisian explanation of transitivity failure requires (Lewis, op. cit., p. 33).

What I am saying, then, is that *the Lewisian* is in no position to use (5) in purported examples of failure of transitivity, since he is committed to denying that (5) is a standard counterfactual. Perhaps it will be retorted that this only shows that there is something wrong with the Lewisian semantics, not that the examples involving (5) are not genuine cases of failure of transitivity. However, I think that the difficulty which (5) presents for the Lewisian semantics should at least make us wary, *whether or not we accept those semantics*, of assuming that (5) *is* a standard counterfactual. It can never be satisfactory to present as a counter-example to an important thesis a case whose interpretation is open to some doubt in this way.

University of Durham,
50 Old Elvet,
Durham DH1 3HN

© E. J. LOWE 1984

COMMENT ON LOWE

By CRISPIN WRIGHT

LOWE agrees ('Wright versus Lewis on the Transitivity of Counterfactuals', ANALYSIS 44.4, October 1984) with my intuitions about the transitivity of counterfactuals, but does not think that I supported them correctly. However, I am not persuaded of his reasons for thinking so. One way of putting my objection to Lewis is that if transitivity is simply invalid for counterfactuals, then there ought in general to be nothing incongruous about a subject's assertion, 'in the same breath', of the premises for a transitivity step in the kinds of case where, Lewis believes, the step fails; incongruity should arise only when the step is, invalidly, made. Lowe allows, as he envisages Lewis allowing, that there often is such an incongruity *before* the transitivity step is made, but suggests that Lewis should have no difficulty in giving a pragmatic (contrast: semantic) explanation of the phenomenon, thereby preempting my suggestion that the real source of the incongruity is the valid entailment, by transitivity, of an incongruous conclusion. It is, however, quite unclear how in general such a pragmatic explanation should proceed. Lowe's specific suggestion (p. 181) would require that it invariably be the case that one of the premisses is 'common knowledge', after the manner of premiss (2) in the snow example. There is no reason why that should be so. Suppose that I, whom you regard as an expert on the Soviet political system and the kind of characters who succeed within it, inform you that 'If Benn had been born and brought up in Russia, he would have failed to achieve political eminence', and that you happen to believe for sound, though not generally known, reasons of your own, that 'if Benn had failed to achieve political eminence, he would have been a market gardener in Lowestoft'. Manifestly, you are not at liberty to conjoin and file, as it were, the two propostions, without qualification. And the explanation of this cannot be the pragmatic one envisaged by Lowe, since (i) neither is common knowledge, and (ii) no simultaneous assertion of both is anyway being made.

Accordingly, I stick to it that the right explanation of such cases is that the transitivity step is valid, but leads to an unacceptable consequence; and, consequently, that no 'filing' may reasonably take place until the premises have been appropriately reformulated or the grounds for their assertion reassessed. In any case, there are not the 'dire implications' which Lowe finds in my proposal. The proposal is not to relativize the *validity of inferences* to context but to relativize the *content* of counterfactuals to context. What would indeed be dire would be if one and the same proposition, identified across different contexts, possessed only context-dependent logical liaisons. But that is not in view. What I meant to suggest was, rather, that the right account of counterfactuals – and indeed of conditionals in general – may very well involve attributing to them something like, though not to be identified with, token reflexivity: their truth conditions, that is to say, will be a function both of a context-independent sense and of an envisaged context of utterance. The relevant notion of 'context' will involve, of course, not merely aspects of the occasion of utterance but considerations to do with the informational presuppositions and purposes of speaker and audience. How the analysis of the notion should best be approached, and whether in particular Lewisian semantics have a part to play, is another question.

A final point about Nozick. Suppose we accept that, as suggested in my previous discussion ('Keeping track of Nozick', ANALYSIS 43.3, June 1983, p. 138), a convention of the following sort operates:

(C) When a number of counterfactual conditionals are in play in a single context, some *single* range of relevant worlds governs the assessment of them all.

For Nozick, it is essential that both the following counterfactual conditionals can be true.

(i) If I were a brain in a vat, I would have no hand;

and

(ii) If I had no hand, I would not believe that I had a hand.

Plausibly, (i) holds irrespective of the range of worlds in view, since it articulates an analytic implication of what it is to be a brain in a vat; while the possibility of the truth of (ii) is exactly what, on Nozick's account, saves the possibility of my knowledge that I have a hand even if we concede the unassailability of the sceptical doubt whether I am a brain in a vat. Now transitivity will, of course, yield an unacceptable consequence from (i) and (ii). But C can be harnessed to generate trouble for Nozick by a different route.[1] For, in the presence of that principle, the validity of the following pattern of inference follows:

$$\frac{P \mathrel{\Box\!\!\rightarrow} Q;\ Q \mathrel{\Box\!\!\rightarrow} R}{P \,\&\, Q \mathrel{\Box\!\!\rightarrow} R}$$

whereby (i) and (ii) can be made to yield the unacceptable

(iii) If I had been a brain in a vat and had had no hand, then I would not have believed that I had a hand.

What follows is that Nozick must deny C or admit that, the analyticity of (i) notwithstanding, (i) and (ii) may not be maintained in a single context. And, of course, the cost of the latter option is that (ii) may not be maintained in the context of dialogue with the sceptic.

Department of Logic and Metaphysics,
The University, St Andrews, Fife KY16 9AL

© CRISPIN WRIGHT 1984

---

[1] Stig Rasmussen persuaded me that this might be a helpful way of presenting the point.

INUS CONDITIONS

By A. J. DALE

IN his recent paper 'On the Nature of *INUS* Conditionality' (ANALYSIS 44.2, March 1984, pp. 49–52) T. C. Denise claims to have repaired a flaw discovered by F. Jackson in Mackie's definition of an *INUS* condition. I shall argue that Mackie's definition does not have the purported flaw, that Denise's modification would not in any case succeed in repairing the flaw, and that the modification runs counter to Mackie's explicit intentions and makes the definition useless as part of an analysis of 'A caused P'.

The definition which Denise quotes from Mackie is as follows:

> 'A is an *INUS* condition of a result P if and only if, for some X and for some Y, (AX *or* Y) is a necessary and sufficient condition of P, but A is not a sufficient condition of P and X is not a sufficient condition of P' (Mackie, 'Clauses and Conditions', *American Philosophical Quarterly* 2, 1965, pp. 245–64).

The flaw in the definition, Denise claims, is revealed in F. Jackson's review of Mackie's paper (*Journal of Symbolic Logic* 47, 1982, pp. 470–3). It is that *any* irrelevant condition will turn out to be an *INUS* condition under that definition. Consider conditions S_1 and S_2 such that S_1 *or* S_2 is a necessary and sufficient condition for P. Then for any A such that neither A nor $\sim A$ is sufficient condition for P or S_1

(i) $A.(\sim A \text{ or } S_1) \text{ or } (S_2 \text{ or } \sim A.S_1)$

is a necessary and sufficient condition for P where neither A nor ($\sim A$ *or* S_1) is a sufficient condition for P. Thus, it is claimed, A is an *INUS* condition for P under Mackie's definition. But this is too quick: it ignores Mackie's explicit qualifications. He amplifies his definition in the paragraph following his definition in order to forestall misunderstandings. He explicitly states that the disjunction should consist of *minimal sufficient conditions*. Jackson's putative counterexample ignores this requirement since it contains the disjunct $\sim A.S_1$ which is not a minimum sufficient condition since S_1 is *ex hypothesi* sufficient alone. Denise's further discussion of Mackie's article 'Mill's Methods of Induction' (*The Encyclopaedia of Philosophy*, ed. P. Edwards, 1967, vol. 5, pp. 324–32) shows a total disregard of the proviso that the Y of the definition should be a disjunction of minimum sufficient conditions. For, though Mackie does explicitly use in illustrating a variation on one of Mill's methods a case where ($A.B.\sim C$ *or* $A.\sim B.C$ *or* $\sim A.B.\sim C$) is a necessary and sufficient condition for P, it is not the case that this justifies Denise's assertion that under the definition A and $\sim A$ are here *INUS* conditions for P. In this example $B.\sim C$ is itself a sufficient condition for P and so A does not fulfil the requirements of an *INUS* condition since in the definition X must not be sufficient for P nor must Y contain any but minimum sufficient conditions.

However, there are other examples which cannot be pushed aside quite so summarily. Suppose S is a necessary and sufficient condition for P and neither A nor ${\sim}A$ are sufficient for P or S. Then:

(ii) $A.(S \textit{ or } {\sim}A) \textit{ or } {\sim}A.(S \textit{ or } A)$

is necessary and sufficient for P where neither A nor $(S \textit{ or } {\sim}A)$ is a sufficient condition for P and ${\sim}A.(S \textit{ or } A)$ does not contain any conjuncts which are themselves sufficient for P. This example is an extension of J. Kim's which showed that almost any condition can be incorporated as a conjunct in a minimum sufficient condition (J. Kim, 'Causes and Events: Mackie on Causation', *Journal of Philosophy* 68, 1971, p. 433). As Kim notes Mackie indicates that only sentences in disjunctive normal form should be considered in the formulation of the necessary and sufficient conditions for P. (Mackie, 'Causes and Conditions', p. 255). Now limitation to formulations in disjunctive normal form has its own disadvantages in this context (*vide* Kim) but it does free the definition from the accusation that it allows irrelevant conditions to be *INUS* conditions, for (ii) is logically equivalent to the normal form $A.S \textit{ or } {\sim}A.S$ and the sufficiency of S guarantees that this formulation fails to give A *INUS* status. It will thus be impossible to find an example in which some irrelevant condition occurs as A in $(AX \textit{ or } Y)$ where $(AX \textit{ or } Y)$ is in disjunctive normal form and otherwise satisfies Mackie's conditions. Each of the examples of irrelevant conditions purportedly deemed an *INUS* condition under the definition trades on ${\sim}A$ occurring as a disjunct in X. Thus $A.{\sim}A$ would occur in its expansion and would be eliminated by simplification to a disjunctive normal form.[1]

Denise proposes to repair Mackie's definition by inserting an extra clause, *viz*: A should be a necessary condition of $P.X$. This addition entails that the A of (i) will not be an *INUS* condition, for if S_1 and ${\sim}A$ then both P and X obtain but A does not obtain. Similarly for my example (ii), if S and ${\sim}A$ then both P and X obtain but A does not.

Unfortunately Denise's modified definition is subject to the same objection: it grants *INUS* status on irrelevant conditions. Suppose S is a necessary and sufficient condition for P; then so is

(iii) $A.({\sim}S.\ {\sim}A \textit{ or } S.\ {\sim}S) \textit{ or } S$

If neither A nor ${\sim}A$ is sufficient for P or S then A is an *INUS* condition for P under Denise's revised definition since if P obtains then so does S. Now S together with ($\sim S . \sim A$ *or* $S . \sim S$) entails A (or any other sentence for that matter!) and so A is a necessary condition of $P.X$ where X here has the form ($\sim S.\ \sim A$ *or* $S.\ \sim S$).

---

[1] Since it is not altogether clear in the literature what counts as a formula in disjunctive normal form I should make it clear that I use the term to exclude any conjunction (within the disjunction) which contains a sentence and its negation. Otherwise Mackie's definition is open to the trivial $A.\ {\sim}A \textit{ or } S$ as a necessary and sufficient condition for P.

Once again a formulation in disjunctive normal form disposes of A as a candidate for *INUS* status as (iii) reduces to S.

Moreover Denise's proposal runs counter to Mackie's explicit intentions for Mackie makes allowances for *different* A and B to be conjuncts in two minimal sufficient conditions whose other conjuncts are identical, i.e. both $A.X$ and $B.X$ may occur in the disjunction (Mackie, ibid., p. 248). It is also far too strict a condition when the purpose of Mackie's introduction of *INUS* conditions is recalled. Mackie maintains that a statement of the form 'A caused P' often makes implicitly the claim that A is (at least) an *INUS* condition for P. Now Denise's addition to the definition would mean that we would by making such a statement be implicitly asserted that $\sim A$ is a necessary condition for each of the other minimum sufficient conditions. We should then be prevented from correctly claiming that A caused P unless $\sim A$ was so incorporated. But this consequence is surely undesirable. Consider a switching circuit composed of switches A′ and B′ in series and these in parallel to C′ and D′ in series. If A is the condition that A′ is closed etc. then a necessary and sufficient condition for the current to flow is:

$A.B$ *or* $C.D.$

Under Mackie's definition but not under Denise's modification A is an *INUS* condition for the current to flow, for it is perfectly possible for the current to flow and B to obtain without A obtaining – precisely when C and D obtain. If in fact the circuit was in the state of B′ closed and C′ open there is no doubt that the closing of A′ would be deemed to be the cause of the current's flowing as indeed it would on Mackie's definition since he requires in addition to A's being an *INUS* condition, that it should also obtain, that B should obtain and that $C.D$ should not. Denise, on the other hand, does not allow that A is an *INUS* condition and so it could not in this situation cause the current flow. This alone is sufficient to rule out the proposed modification to Mackie's definition. There is much that is wrong with Mackie's definition as Kim's cited paper has made clear, but Denise's modification does nothing to cure those particular ills.

University of Hull,
Hull HU6 7RX

© A. J. DALE 1984

GOLDSTEIN ON QUOTATION

By D. K. BUCKNER

LAURENCE Goldstein ('Quotation of Types and Other Types of Quotation'. ANALYSIS 44.1, January 1984, p. 1) disputes the commonly held view that the sign formed by placing quotation marks around another sign is a sign for the other sign. He says (p. 4) 'One can think of . . . quotation marks as arrows pointing to a sample which could be remote from the sentence'. Davidson ('Quotation', *Theory and Decision* 11, 1979, pp. 27–40, reprinted in *Inquiries into Truth and Interpretation*, Oxford 1984, pp. 79–92) holds a similar view: a 'demonstrative' account of quotation. The account seems to me quite implausible, since it requires, ultimately, that we cannot read conventionally quoted expressions in the way that we do.

Consider:

(1) This mark ☺ resembles a face.

When I read this, I treat the gap formed by the face like the gap between, e.g., the last word of p. 1,033 and the first word of p. 1,034. I discontinue reading at 'mark', and resume it at 'resembles'. If I pause and look at the face, I treat this as a mere perceptual interlude, the sort of thing that might happen if I tried to separate p. 1,034 from p. 1,035, if they were stuck together. (The same thing happens when a narrative is interrupted by coughing or sneezing, or any other noisy event. The interrupting noise is not treated as a grammatical part of the narrative. The professor is interrupted by a gigantic explosion as he reads a scholarly paper. He looks up, shrugs, and continues the paper exactly where he left off.)

In this respect (1) is exactly similar to

(2) This word: Kuznetzin: denotes the former Russian ambassador.

When we read (2), what we *read* is 'This word denotes the former Russian ambassador'. We do not read 'Kuznetzin', we jump straight from 'word' to 'denotes'; if we do pause and look at 'Kuznetzin' then it is just like looking at or examining the face in (1). We use the pause forced on us by the colons as an opportunity to look at the word inside them, but we do not read that word — at least, *not as a part of the sentence*. It is a sample that we might look at when we read the sentence, but which is not a part of that sentence. As Goldstein says, it could be spatially remote; (2) could be rewritten as

(3) This word denotes the former Russian ambassador.
↓
Kuznetzin

Davidson's account is no different in this respect. He says (p. 38, my emphasis):

> '... taken in abstraction from semantics, the question of location is trivial. In spoken sentences, temporal sequence plays the role of linear arrangement in writing. But if I say "I caught a fish this big" or "I caught this fish today", my hands, or the fish, do not become a part of the language. *We could easily enough remove the quoted material from the heart of the sentence.*'

So it is fundamental to the demonstrative account that the "quoted" sample is not read; and this seems true if the sample is separated from the sentence in the way I have considered (e.g. by double colons, labelled boxes, arrows etc.). But it is not true at all of conventionally quoted material, where we use inverted commas and an appropriate grammatical form. We *do* read the 'Kuznetzin' of

(4) The word 'Kuznetzin' denotes the former Russian ambassador.

and without it the sentence does not have the required sense. If we bracket off the quoted word as in (2), (3), then we must read 'The word denotes the former Russian ambassador', which does not make a great deal of sense, unless we mentally supply a preceding sentence like 'There is a word written on the wall of the Russian embassy'. Remove a word which is *truly* quoted (i.e. that is not bracketed off by colons, arrows, or any other device that creates a parenthesis) and you utterly destroy the intended reading of the sentence. (Try, for instance, to remove the quoted expressions in the third sentence of the Davidson quote. You get the absurd 'But if I say or, my hands, or the fish, do not become a part of the language'.)

That is not to say that quotation marks do not have some role in drawing our attention to the quoted material. But they do not detract from our *reading* that material either. When we read (4), perhaps the reading goes like this: 'The word "Kuznetzin" (pause, look back at the word in quotes) denotes the former Russian ambassador'. *First* we read the quoted material; *then* we look back to examine it, or perhaps cast a sideways glance at it as we read it. Whether or not we *examine* the quoted word, we must first of all read it; the demonstrative account says that we do not read it; and therefore, I claim, it must be wrong.

University of Bristol,
Wills Memorial Building,
Queens Road,
Bristol BS8 1RJ

©D. K. Buckner 1984

EXPLAINING DONNELLAN'S DISTINCTION – A REPLY

By D. E. OVER

IN 'Explaining Donnellan's Distinction' (ANALYSIS 44.1, January 1984, pp. 13–14) Jeffrey King and Michael Liston question my attempt to account for the referential/attributive distinction in my paper, 'Effective and Non-Effective Reference' (ANALYSIS 43.2, March 1983, pp. 85–91). My paper implied that a speaker S uses 'the thirty-seventh prime' (to take the example given by King and Liston) referentially in asserting

(1) The thirty-seventh prime is odd

only if S associates with this definite description an effective method for deciding whether any given number is its referent. King and Liston try to cast doubt on this position by describing an example (on p. 14) in this way:

> Suppose S, himself being slow mathematically and not knowing that all primes greater than two are odd, asserts (1) solely on the basis of having been given the number *a*, which he notes is odd, by Ramanujan, a recognized mathematical genius. Suppose also that in asserting (1), S uses the description 'the thirty-seventh prime' to pick out *a* and to say of it that it is odd. As such, his use is a paradigm of Donnellan's referential use. But though S can now pick out the thirty-seventh prime, in what sense does S have an effective method for picking out the thirty-seventh prime?

Surprisingly, King and Liston do not say how S has been given the number *a* by Ramanujan. I try to stress in my paper (p. 87) the importance of explaining the 'favoured method' (as Dummett calls it) of being given objects of the relevant type before the notion of an effective method is applied. Perhaps King and Liston are presupposing that the favoured method of being given natural numbers is as the referents of numerals in our positional notation. In that case Ramanujan has used '157' to try to tell S which number is the thirty-seventh prime. But '157' is itself a complex term in positional notation, and I hold that S uses it referentially only if he understands an effective method for counting up to 157 (using the other numericals or some other marks or sounds) and then stopping. Without knowledge of such a method, however trivial it may appear to be, S does not really know which number Ramanujan was talking about, and he can at best use '157' attributively.

Let us contrast the above example with one in which Ramanujan proves only non-constructively that there is a number satisfying an undecidable predicate P. Now it follows from my proposal that Ramanujan can only use 'the smallest number with P' attributively, as he is unable to generate effectively a numeral with the same referent, and I think that this is intuitively the correct result. It should also be observed that picking out numbers by counting with numerals would be a favoured method in any reasonable philosophy of mathematics. The process of generating the numerals would be seen either as helping to construct the numbers in some sense (as in formalism and intuitionism) or as giving us our basic intuitions of them (as in various types of realism).

King and Liston also ask how there can be an analogy between effective methods in mathematics and what I claim to be effective methods underlying Donnellan's example of ordinary physical objects. As noted in my paper, Donnellan's examples concern ordinary objects we can observe together in one place, such as people at a party, and I accepted (p. 87) Dummett's point that the favoured method of separating one of these objects from the others would be by using a demonstrative or an indexical. (Under what theory of these objects would this not be the favoured method?) Now there surely is a close analogy between effectively finding the smallest number in a finite set of natural numbers and, say, a method we might use to point out the shortest person at a party. And I cannot see that it really matters whether the methods we use in cases like these are quite trivial. (At one point on p. 14 King and Liston even seem to doubt the existence of trivial recursive functions.) But perhaps we should make our terminology more naturally fit occasions when we exercise simple recognition capacities, as when we see immediately that 2 is a smaller number than 3 or that a toddler is shorter than his father. The general definition would then be that a speaker uses a singular term with effective reference just in case he has the ability to decide (just by looking and seeing or by following some non-trivial effective method) which object is the referent of the term.

King and Liston conclude that I am merely substituting one label for another. But effectivity is not a label I have invented. It is a concept essentially connected with the distinction between constructive and non-constructive reasoning and between decidable and undecidable statements. Effective reference is simply the property possessed by singular terms used in constructive reasoning and present in effectively decidable statements. These constructive concepts were known about and investigated long before Donnellan introduced his distinction, and recently they have received a great deal of attention from Dummett and his followers (though effective reference should be much more explicitly studied). It seems to me to be important in itself to bring out the connection between the referential/attributive distinction and other distinctions of fundamental interest.

The connection I establish also has the positive point of suggesting a way to generalize the referential/attributive distinction. Consider Russell's hierarchy of examples described in a passage at the end of 'Knowledge by Acquaintance and Knowledge by Description' in *The Problems of Philosophy*:

> It will be seen that there are various stages in the removal from acquaintance with particulars: there is Bismarck to people who knew him;

> Bismarck to those who only know of him through history; the man with the iron mask; the longest-lived of men. These are progressively removed from acquaintance with particulars; the first comes as near to acquaintance as is possible in regard to another person; in the second, we shall still be said to know 'who Bismarck was'; in the third, we do not know who was the man in the iron mask, though we know many propositions about him which are not logically deducible from the fact that he wore an iron mask; in the fourth, finally, we know nothing beyond what is logically deducible from the definition of the man.

The people who knew Bismarck could use 'Bismarck' with effective reference. They could point him out as the referent of the name and, through observation, acquire information about him. By doing this they would ground the name, and they could use it informatively in constructive reasoning and justification. Grounding in this sense is a constructive notion: a speaker grounds a term he uses with effective reference just in case he does exercise his ability to decide which object is its referent. At the other extreme is Russell's use of 'the longest-lived of men', which is completely ungrounded. Russell could only use this term in relatively uninformative non-constructive reasoning and justification, e.g., to support the statement that the longest-lived of men is a man. (There are general, non-constructive reasons for thinking that the term has a referent, and then trivial logical reasons for inferring that that referent was a man.) My analysis implies that the people who knew Bismarck could use his name referentially and that Russell could only use 'the longest-lived of men' attributively. But what should be said about *our* use of 'Bismarck' and of 'the longest-lived of men'? In the case of 'Bismarck' we have been passed a grounded name along with much information which could have been (and was) used to point out its referent. By analogy there is good reason for holding that the availability of this information does enable us to decide which historical figure the name refers to. We have that solid basis for saying that we can employ 'Bismarck' in constructive reasoning, using it with effective reference and so referentially. In the other case we have been passed 'the longest-lived of men' by someone who did not ground it in its referent. Having no way of getting any more information about this referent than he had, we can only use this term with non-effective reference and so attributively. Of course, much more needs to be said about grounding, information, and the other constructive concepts, but at least we can now see a promising way forward.[1]

---

[1] Theories of different notions of grounding have been presented by Michael Devitt in *Designation* (Columbia University Press, New York, 1981), and more significantly by Gareth Evans in *The Varieties of Reference* (O.U.P., Oxford, 1982). Kripke's concept of grounding, as found in an 'Outline of a Theory of Truth', *Journal of Philosophy*, 1975, pp. 690–716, does not directly concern singular terms, but is of some indirect relevance here. I try to say more about the relevant constructive concept of grounding in my paper, 'Constructivity and the Referential/Attributive Distinction', which is forthcoming.

Taking this approach to the referential/attributive distinction does lead to some serious problems. From a classical point of view, constructivity is a flexible epistemological concept and not a hard and fast semantic one. Such a view encourages us to think in terms of degrees of effectivity and constructivity, and would allow us to speak of our use of 'the man in the iron mask' as an intermediate case of effective reference. And this point, as I try to indicate in my paper (p. 90), should lead us to follow Kripke and ask whether the referential/attributive distinction is 'exclusive or exhaustive'. The advantage is that this is a question we should ask, and in general we see clearly questions which should be asked about the distinction when we place it in its proper setting, along with constructive concepts and the deep questions to be asked about them. The analysis is a necessary step towards a full explanation of Donnellan's distinction.[2]

Sunderland Polytechnic,
Chester Road,
Sunderland SR1 3SD

©D. E. Over 1984

---

[2] I should like to thank Kit Fine for helpful discussion of the topics in this paper.

NOZICK AND KNOWLEDGE – A REJOINDER

By B. J. Garrett

In 'Knowledge, Reliable Methods and Nozick' (Analysis 44.1, January 1984, pp. 30–3) David Gordon argues that the counterexamples which I produced to Nozick's theory of knowledge ('Nozick on Knowledge', Analysis 43.4, October 1983, pp. 181–4) are unsuccessful, not because they fail to satisfy Nozick's conditions for knowledge, but because they are compelling only if certain premises concerning the conditions necessary for knowledge are assumed, premises for which I do not argue and which a Nozickian will not accept. In this rejoinder I show how Gordon's argument fails.

Note to start with that it is no presupposition of my counterexamples that it is necessary for knowledge that one have good grounds for thinking one's method reliable if it is reliable. For the methods employed by X and Z are *unreliable* given, as I of course intended, that the reference classes by which the reliability of the methods is to be judged are taken to be, respectively, the class of all pairs of persons and the class of all quantified mathematical statements to which no counterexamples have been discovered.

(Note, incidentally, that the example about Z can easily be modified to avoid Gordon's specific criticisms of it. For it is not essential to the point of the example that Z knows that no counterexamples have been discovered to Fermat's Last Theorem. Suppose, instead, that he believes the theorem to be true solely on the grounds that it is not falsified for x = 1, y = 2, z = 3 and n = 4. On Nozick's theory Z knows, via this method, that Fermat's Last Theorem is true if it is true. And this is absurd.)

As to Y, the method he employs (the reference class of which is the class of all even numbers), though assumed to be reliable, is not known (by anyone) to be reliable. Thus the explanation of why Y's belief does not constitute knowledge need not consist in the fact that Y does not know or have good grounds for thinking his method to be reliable but simply in the fact that the method is not known to be reliable. (Hence whether a true belief reached via a reliable method which the believer does not know to be reliable constitutes knowledge will sometimes depend upon whether the method in question is known to be reliable. This fact allows one consistently to hold, for example, both that Y does not know that the number of people in Paris is the sum of two odd primes and that the examples of Alan White ('Doubting One's Methods', ANALYSIS 43.3, June 1983, pp. 133–4) are genuine cases of knowledge.)

However, Gordon is willing to accept that my examples demonstrate that it is possible for an unjustified true belief reached via an unreliable method to conform with conditions (3) and (4) of Nozick's analysis but claims that, even so, my examples will be decisive against Nozick only if justification is necessary for knowledge. And since I profered no argument in favour of such a condition, I have not produced clear counterexamples to Nozick who can thus accommodate my examples as genuine cases of knowledge.

Now certainly it is true that *if* justification is necessary for knowledge then my examples are counterexamples to Nozick. But it does not follow from this, nor is it true, that my examples are counterexamples *only if* justification is necessary for knowledge. For one could perfectly well accept my counterexamples and yet *either* be agnostic about whether justification is a necessary condition for knowledge *or* hold that justification is not everywhere a condition for knowledge. For example, suppose that one knows that p and believes that p via a method which one does not have good grounds for thinking to be reliable (or has good grounds for thinking to be unreliable) though it is reliable and known to be reliable. Then one knows that p even though one is not justified in believing that p. (For example, suppose, modifying one of White's examples, that the man who knows that 1234567896 is divisible by 4, but has good grounds for thinking his method to be unreliable, also happens to believe it. Then he knows this arithmetical truth but is not justified in believing it to be true.)

And one *could* consistently accept that such examples show that justification is not necessary for knowledge and that my examples are counterexamples to Nozick.

Furthermore, even though one may want to explain why my examples are counterexamples by reference to the presence of unjustified beliefs it does not follow from this that a belief can constitute knowledge only if it is justified. For just as one could perfectly consistently explain why, for example, Gettier's examples are counterexamples to the 'justified true belief' theory of knowledge by reference to the lack of a non-deviant causal chain linking the belief that p with the fact that p *without* being forced to conclude that knowledge must have a causal condition as part of its analysis, so one could explain why my examples are counterexamples to Nozick by reference to the presence of unjustified beliefs *without* being forced to conclude that justification is essential to knowledge. Hence my counterexamples are successful even if justification is not necessary for knowledge. Of course, it may well be that, in the final analysis, my examples are counterexamples to Nozick precisely *because* justification is necessary for knowledge. But the point is that, contra Gordon, it is no essential precondition of the success of my counterexamples that this be so.

Not only are Gordon's criticisms off-beam but his *way* of criticising my paper is also defective. For whether or not my examples are counterexamples to Nozick's theory depends, not upon whether some particular condition is necessary for knowledge, but simply upon what we intuitively and pre-theoretically want to say about the deliverances of that theory in the situations which I have described. And any theory of knowledge must be sensitive to our judgements about such hypothetical situations since those judgements reveal the logical shape, and hence the identity, of our concept of knowledge. But nothing in Gordon's reply gives reason to doubt the soundness of the judgement that, in my examples, the relevant statements are not known and hence their status as counterexamples to Nozick has not been impugned.

Christ Church, Oxford ©B. J. GARRETT 1984

A NOTE ON A RESPONSE OF HORNSBY'S

By E. J. LOWE

IN my 'Reply to Hornsby on actions' (ANALYSIS 43.3, June 1983) I claimed that, at least in the case where *A* is an inanimate object, it is absurd to assert something of the form '*A*'s causing *x* caused *x*', where *x* is an event. As an example I adverted to the case in which billiard ball *B* caused billiard ball *C* to move, i.e., caused the event of *C*'s moving, and I pointed out that if we wanted to refer to an *event* which caused *C*'s moving we might well refer to *B's moving,* but certainly not to *B's causing C's moving*, which I suggested is doubtfully an *event* at all. So while we might want to say that *B*'s moving caused *C*'s moving, we certainly shouldn't want to say that *B*'s causing *C*'s moving caused *C*'s moving.

In her rejoinder ('Events that are causings: a response to Lowe', ANALYSIS 43.3, June 1983) Hornsby somewhat misrepresents what I say. She writes: 'I think [Lowe] has made it clear that if we must reject the assumption that actions are events, then so also must we reject the claim that ball *B*'s *hitting* ball *C* is an event' (p. 142, my emphasis). But what I suggested was not an event was *B's causing C's moving*, not *B's hitting C*. On what basis does Hornsby regard these two descriptions as having the same referent? Surely not on the grounds that '*B* hit *C*' is logically equivalent to '*B* caused *C* to move', for there is of course no such equivalence. Perhaps then on the grounds that the referents of these two descriptions are events with the same causes and effects? But this would simply beg the question against me. For my part, I am perfectly happy to regard *B*'s hitting *C* as an event: it is the event of *B*'s coming into contact with *C*. I am moreover happy to accept *B*'s hitting *C* as a cause of *C*'s moving. None of this is inconsistent with my denial that *B*'s causing *C*'s moving was a cause of *C*'s moving, nor with my suggestion that *B*'s causing *C*'s moving is not an *event.* And, for what it is worth, I suspect that common sense is on my side.

University of Durham,
50 Old Elvet, Durham DH1 3HN

© E. J. LOWE 1984

McGINN ON BENACERRAF

By GRAHAM SPINKS

WHILE I found Marie McGinn's discussion of Wright's appeal to Quine's remarks on indeterminacy incisive ('Wright's Reply to Benacerraf', ANALYSIS 44.2, March 1984) I was puzzled by her conclusion. In particular I was surprised by her assertion that Benacerraf had failed to demonstrate that numbers are not sets.

If the claim that numbers are sets is to be understood as the claim that numerals refer to or denote sets then it must be plain that, in showing that every recursive progression of objects has an equal right to be construed as the number series, Benacerraf did demonstrate not only that numbers are not identifiable with sets but that they are not identifiable with any other objects. The principal lesson to be drawn from the identification of the indeterminacy to which Benacerraf has drawn our attention is, surely, that numerals cannot be construed referentially.

It would appear, however, that Dr McGinn's assertion is founded upon the claim that in order to do number-theory it is essential to have committed ourselves to the existence of a recursive progression of objects and that she understands the claim that numbers are sets as the claim that a commitment to the existence of sets is a necessary condition of the practice of number-theory. Such an understanding would have the advantage of permitting one to say that numbers are sets while accepting the indeterminacy of questions about which sets they are. However, there is a gap between the claim that we must commit ourselves to the existence of a recursive progression and the claim that the progression must be a progression of sets; after all, sets are not the only kind of object which can constitute recursive progressions.

Thus while Benacerraf has not shown us how to escape commitment to abstract objects it is not the case that he has 'made it clear that the abstract objects we must commit ourselves to are sets'.

Further, the ontologically parsimonious, if not the anti-platonistic, might find some comfort in Benacerraf's observation that numerical series themselves are recursive progressions, since from this it is clear that we need not commit ourselves to the existence of any progression of objects (sets or whatever) in *addition* to the numerical series. As Benacerraf urges (*Philosophical Review*, 74, 1965, p. 72), there are not two kinds of objects, number-words and numbers, but only the former.

Darwin College,
Cambridge CB3 9EU

© GRAHAM SPINKS 1984

NECESSARY AGNOSTICISM?

By ROBERT MCLAUGHLIN

'A THREE-HEADED hippogriff is alive and well right now on the fifth planet of the Vega system.' If I followed this claim with the frank admission that I had *no* grounds at all for it, you might be inclined to reject it out of hand, despite your presumed lack of any specific grounds for its negation. To do so would be to invoke a version of Occam's Razor: 'total lack of grounds for an ontically significant proposition is itself a sufficient ground for its negation'.[1] With respect to an ontic proposition in a justificatory vacuum, the Razor would rule out agnosticism (suspension of judgement) and enjoin scepticism. This parsimony principle ordinarily has the status of a canon of rationality.

Is the Razor ever *inapplicable* to gratuitous ontic claims? Pottinger,[2] a Friend of the Ontological Argument, apparently thinks so. Having satisfied himself that God's existence is either necessary or impossible, he concludes that 'the atheist has a very tough row to hoe' (op. cit., p. 46). This suggests that we atheists have a much harder task in a modal setting: we must show, not merely that 'God exists' is totally unjustified, but further that its negation is *necessary*. Suppose that all non-modal arguments for God's existence fail; and that proponents[3] of the modal (ontological) argument are unable to justify 'God is possible' (whence, if they could justify it, given the purported modal character of 'God exists', they could infer 'God is necessary'). Even then, Pottinger seems to say, this justificatory impotence on the part of theists would fall far short of a case for atheism. If atheists cannot directly prove that God is impossible, he seems to urge the necessity of *agnosticism*, as entailed by the premiss of a modal justificatory vacuum plus the necessary (?) principle that *assent to a necessary claim (or its negation) must be withheld in the absence of modal justification*. Patently the agnostic conclusion here requires that Occam's Razor does not apply to gratuitous *modal* ontic claims. If it did, then equally gratuitous claims for the existence of a (necessary or impossible) God or a (contingent) hippogriff would warrant the same response: rejection. I shall argue that it does, and so they do.

The notion that *only* modal justifications can be relevant to the appraisal of modal propositions can be refuted readily by cases, of which the history of mathematics is a convenient source. A modal proposition like *Q*: 'there exists no general solution of the quartic equation' can be *disproved* simply enough, by counterexample – namely, by producing a general solution algorithm for the quartic.[4] Disproof by counterexample does not, of itself, amount to a demonstration of the *necessity* of the claim's negation; even so, it is a sufficient disproof, a fully adequate ground for rejecting the claim. When its demonstrated falsity is then conjoined with its acknowledged modality, the claim is seen to be necessarily false; but a non-modal refutation of the claim can quite properly precede this revelation of necessary falsity. Direct modal demonstration of the necessary falsity of *Q* – e.g. from Galois' group theory (see Eves ibid., pp. 378–80) – also may turn up; but, as we see, such modal demonstrations are in no way essential to the appraisal of modal claims.

---

[1] This is milder than the traditional form: 'Entities are not to be multiplied beyond necessity' (Lacey, A. R., *A Dictionary of Philosophy*, London: R. & K. P., 1976, p. 147). Ironically, the latter would appear to enjoin scepticism in case theists failed to demonstrate the *necessity* of God! The more reasonable version of the Razor invoked in the present discussion requires only that *some* grounds, even non-modal ones, be available, in order for an ontic claim to avoid rejection. Atheism is warranted by this Razor only in a *total* justificatory vacuum for theism.

[2] Pottinger, G. 'A Formal Analysis of the Ontological Argument', *American Philosophical Quarterly 20*(1): 37-46 (1983). Pottinger develops his account within the modal predicate calculus *free S5*. As he notes, there are *various* formal modal systems, exemplified by Hughes, G. E. & Cresswell, M. J., *An Introduction to Modal Logic*, London: Methuen, 1968, for instance. The simple epistemic points I make here are independent of differences among such systems, relying only on common elementary intutitions about modality (ibid., Chapter Two). I call proposition *modal* iff either it or its negation is necessary; and by a *modal justification* of a proposition I mean a demonstration of its necessity — typically by deduction from necessary premisses.

[3] As well as Pottinger (op. cit.) see Plantinga, A., *The Nature of Necessity*, Oxford: O.U.P., 1974 — esp. Chapter X; Hartshorne, C., *The Logic of Perfection*, La Salle, Illinois: Open Court, 1962 — esp. Chapter Two; Malcolm, N., 'Anselm's Ontological Arguments', *The Philosophical Review LXIX*: 41-62 (1960). The latter two do *try* to justify 'God is possible'. The case that concerns me is where such attempts *fail*, if made at all: does such justificatory failure or lack rationally warrant atheism, or merely agnosticism?

This insight can be taken further: as well as deductive grounds like counterexamples, non-deductive grounds also may be relevant to the appraisal of modal claims. Thus a 16th Century algebraist (say, Lodovico Ferrari: see Eves ibid., p. 220) might well have been led to judge *Q* false, on the *inductive* grounds that general solutions had already been found for the quadratic and cubic equations, so that by extrapolation a general solution for the quartic might be expected. Such grounds, admittedly non-demonstrative, are nonetheless not irrelevant to judgements of the truth-value of the claim. So, that a claim is modal does not preclude *divers* non-modal deductive and even non-deductive grounds from bearing on its appraisal.

The notion that modal claims are exempt from non-modal appraisal may spring from conflation of the *epistemic* question, whether a claim is justified, with the *logical* question, whether it is modal. So conflated, a claim's (purported) modal character is taken to give it a privileged epistemic status, which would exclude non-modal appraisal grounds. This conflation in turn may arise from a blurring of the relation between *necessity* and *modality*. To call a claim necessary is indeed to imply that it is justified, *and* that no other considerations can affect its justification. To dub a claim modal is to imply that *if* true it is necessarily true, which of course is *not* to say that it is justified; nor – and this is crucial – is it to say that no rational considerations other than modal ones can bear upon its justification.

Occam's Razor is a non-modal rational appraisal principle which applies to ontic claims in justificatory vacua. In the light of the foregoing discussion, the Razor might be expected to function in modal, as well as non-modal, contexts. But if agnosticism about unjustified modal claims is to be warranted, this must be because the Razor, for some reason, is not applicable thereto. Is there any basis for supposing such specific inapplicability of the Razor in modal settings? A *prima facie* case against use of the Razor here has been suggested to me by Keith Lehrer: some modal claims, in cases of justificatory impotence, do seem to be exempt from rejection by the Razor; and if some, why not all? Examples of modal ontic claims as to which agnosticism rather than scepticism might seem rational include 'sets exist' or 'the number seven exists'. These famous claims, of course, are not lacking in (candidate) justificatory grounds (nor are their negations), and so it is easy enough to see the Razor's impertinence to them; for it functions only in the absence of *any* justification. Agnosticism about such claims presumably is warranted, if at all, because the rational grounds for and against them are judged to be roughly balanced, rather than because the claims themselves are somehow immune to the Razor. Nevertheless, there *are* justificatory vacua in which the Razor seems not to apply – such as the case of the mathematical proposition, 'there exists a greatest prime-pair'. At present this proposition has been neither proved nor disproved (in contrast to 'there exists a greatest prime'); yet presumably no one would be tempted to *deny* this claim on the strength of Occam's Razor alone: without an argument either way, mathematicians currently *suspend judgement* on it, and rationally so. Since this and similar claims are distinctively modal, and furthermore lack any present justification, it might seem plausible to suppose that the Razor is specifically inapplicable to gratuitous modal ontic claims.

---

[4] Such an algorithm was published by Girolamo Cardano in his *Ars Magna*, Nuremberg, 1545, drawing on work of his pupil, Lodovico Ferrari. See Eves, H., *An Introduction to the History of Mathematics* (Fourth Edition), New York: Holt, Rinehart & Winston, 1976, pp. 219–23.

Now this challenge to the use of the Razor relies on conflation of the *formal* with the *modal* character of some existentially quantified locutions. Consider the following.

A: '(E!x)(x is the number seven)'.
B: '(E!x)(x is the greatest prime-pair)'.

Despite their grammatical similarity, *A* and *B* may express categorially different propositions. *A* resides in the context of *metaphysical* debates over Platonic realism with respect to numbers and other candidate abstract entities, like sets and properties; it is a strong ontic thesis. While *B* might admit of a similar metaphysical construal, in that the greatest prime-pair may be asserted to exist Platonically, its more natural place is in the setting of *mathematical* discussions of number theory, where it is tantamount to the thesis: 'The sequence of prime-pairs is finite.' So construed, *B* lacks any special ontic import, and its quantifier '(E!x)' has a formal rather than ontic role. Appreciation of this contrast answers the above challenge to the use of the Razor. Insofar as some claims are formal, like *B*, they are not subject to the Razor, which applies only to ontic propositons; so that, in a justificatory vacuum, agnosticism respecting them is rational. This immunity from the Razor obtains just because formal claims are not ontically significant, rather than because they are modal. In contrast, *A* is ontically significant, and so, if there were no grounds for it, its rejection would be enjoined by the Razor.

The claim 'God exists' is patently of ontic significance. So it is a candidate for the Razor; and if it entirely lacks justification, then, modal or not, the Razor justifies its rejection.

Friends of the Ontological Argument have a tendency to obfuscate this otherwise obvious insight: typically they reach a stage in their presentations where they point to a *possibility* claim, namely that the bearer of some specified properties in superlative degree is a possible entity, as a premiss. They then suggest that atheists are somehow obliged to refute this possibility claim — that is, to prove the impossibility of the entity in question — in order to refute theism. Failure at this, given that theists also fail to prove their case (i.e. to show that God is possible), is taken to compel agnosticism. Against this strategy, I have urged that no such heavy *onus* weighs upon atheists. Pottinger *et al.* have put the modal cart before the epistemic horse. Modal claims can be rationally rejected without first proving the *necessity* of their negations; rather, as noted above, the latter will logically ensue from their rational denial. Allowing (which is moot) that 'God exists' is modal, all that atheists need do is justify, by *any* rational means, its negation; from which pair of premisses (that it is modal, and that it is false) it will follow that its negation is necessary, i.e. that God is impossible. The judgement that 'God exists' is false (thence necessarily false) will not be itself a *necessary* judgement, if warranted by non-modal rational grounds; it will be, nonetheless, a rationally justified one. On the starting supposition that there is no ground at all, other than its alleged modality, in favour of 'God exists', the Razor is wholly adequate rationally to warrant its *rejection* — i.e. to justify atheism, and so to refute agnosticism as well as theism. Another way of putting this is to remark that in the (purportedly) unique case of God, whose possible existence implies his necessary existence, the possibility claim itself has (extreme) ontic import; and so, if gratuitous, is to be rejected by the Razor — whence agnosticism, with theism, is refuted.

Of course, the Razor itself needs justification, which I think can be furnished through a strategy of pragmatic vindication; but that is another tale.[5]

Macquarie University,
North Ryde, New South Wales 2113, Australia

© ROBERT MCLAUGHLIN 1984

---

[5] I am grateful to Keith Lehrer and Doug Busch for stimulating the ideas developed here; they may well disown my conclusions.

CAUSE IN PERCEPTION: A NOTE ON SEARLE'S *INTENTIONALITY*

By GREGORY MCCULLOCH

SUPPOSE that

(1) Beryl sees a pig.

It is plausible to suggest that (1) can only be true if Beryl has a visual experience which is caused (in an appropriate way) by a certain pig (the one she sees). It is also at least quite sensible to hold that (1) is a derivative locution that needs to be analysed in terms of

(2) Beryl sees that a pig is present,

where (2) reports that Beryl is in an intentional state with a propositional content specified by the sentence 'a pig is present'. Searle subscribes to both of these views (J. R. Searle, *Intentionality*, Cambridge 1983, pp. 40ff., 47ff.), which in this short paper I do not question. I am concerned with a third view, linked to them, which Searle tries to establish. This is the surprising view that the alleged necessity for there to be a causal relation between Beryl and a pig, if (1) is to be true, is explicitly determined by the propositional content ascribed to her in (2). Searle's view is that the truth of (2) (and hence that of (1)) requires Beryl to have a visual experience whose content is specified by 'a pig is present', and that this content itself involves the causal relation in the sense that, when the relation holds, it is somehow directly known by Beryl (p. 48ff.). Now this is obviously an exciting thesis, particularly in view of traditional Humean scepticism about cause that takes off from the thought that causal relations themselves are never observed (Searle takes his argument to constitute a partial rebuttal of such scepticism: see ch. 4). I like the conclusion, even whilst finding it hard to believe: but there is something wrong with Searle's argument for it. The argument is contained in the following passage:

> The intentional content of the visual experience [had by Beryl when (2) is true] determines under what conditions it is satisfied or not satisfied ... Well, what must be the case ... *in order that the experience be a veridical one?* At least this much: the world must be as it visually seems to [be], and ... its being that way must be what causes [her] to have [that experience]. (pp. 48-9; my emphasis)

The flaw lies in the phrase that I have emphasised, where Searle identifies the satisfaction conditions of the content of a visual experience with those conditions that ensure that a veridical experience is had (cp. pp. 39-43).

Searle's mention of 'satisfaction conditions' picks up his analysis of intentional states. According to this analysis, one learns about the nature of these states by attending to the details of sentences that are typically used to report their presence. Searle's more or less orthodox view is that such reports – (2) is an example – exhibit a *subject-mode-content* structure, which provides a clue to the nature of intentional states themselves. Searle focuses on the *content* component, which is introduced by the sentence ' a pig is present' in (2), claiming that an analysis of content is part of the elucidation of intentionality in general. He holds that content is to be glossed in terms of the *satisfaction conditions* of content-introducing sentences like 'a pig is present'. Satisfaction conditions are states of affairs in 'an independently existing reality' which are somehow determined by the content of the sentence in question, and which must obtain if the sentence is to be satisfied (see p. 7ff.). One job for a theory of content is to say what satisfaction conditions the content of some given sentence determines.

In the context of this sort of account, one can ask at least two questions about a report like (2):

(i) what are the satisfaction conditions determined by the content-specifying sentences?
(ii) is the report true?

That these are distinct questions is clear enough. For an answer to (i) we should look to the works of Frege or some similar authority on semantics and meaning. For an answer to (ii) we should look to how things are with Beryl and her surroundings at the appropriate time. Question (i) is concerned with *what* Beryl sees if she sees at all. Question (ii) is concerned with *whether* she sees.

Searle's question 'What must be the case in order that the experience be a veridical one?' is intended to be a version of (i), since he is enquiring into satisfaction conditions. But his reply seems subtly to run together answers to questions (i) *and* (ii). Beryl cannot have a veridical experience unless (a) things are as they seem and (b) the seeming is appropriately causally related to how things are: this may be true. But whereas condition (a) is the sort of thing that one elucidates by asking question (i), condition (b) seems to belong with question (ii). Beryl just will not *see a pig* – i.e. (2) will not be true – unless condition (b) obtains. But she may well have an hallucination with exactly the same content; then it would *seem* that a pig is present.

Of course, Searle could reply that the content of this hallucination would also (falsely) determine that condition (b) obtain. So it might, if Searle's conclusion is true. But the point is that his argument does not establish this. As long as it is granted, as it must be, that Searle's answer to his question brings in matters that are *prima facie* nothing to do with it, then obviously further defence is required: defence that is offered neither in the passage cited, nor elsewhere in the book.[1]

The University,
Leicester LE1 7RH

© GREGORY MCCULLOCH 1984

---

[1] I am grateful to Professor David Hamlyn and the Editor for criticisms of earlier versions of this paper.

A PARADOXICAL TRAIN OF THOUGHT

By MARTIN HOLLIS

A THINKS of a number and whispers it privately to *C*. *B* does the same. *C* tells them, 'You have each thought of a different positive whole number. Neither of you can work out whose is the greater'. The party breaks up and they go their separate ways.

Sitting alone in his homebound train, *A* muses as follows. 'I picked *157* and have no idea what *B* picked. So, assuming that he indeed chose a different positive whole number, *C* is right. I wonder what follows from that. Well, clearly *B* did not choose *1*, as he would then be able to work out that mine is greater; and by the same token he knows that I did not choose *1*. So he did not choose *2*, since he could then use the previous reasoning to prove that my number is greater. Similarly he can know that I did not choose *2* either. With *2* out of the way, I infer that he did not choose *3*; and he can infer that I did not choose *3*'.

The train begins to pick up speed. 'So he did not choose *4*; nor, he infers, did I. *5* is out in the same way. So is the next number. The mathematical induction is well grounded by now. Indeed it is unstoppable. I can keep this up for ever. But that is absurd. It means that I cannot have picked *157*, which I certainly did.'

The engine whistles derisively. 'Wait. Perhaps it follows only that *C* was wrong. What about? Well, not that we chose different numbers. It can't possibly follow that *B* chose *157* too; and, anyway, the same trouble occurs, if we had drawn our numbers from a hat or agreed that I would choose an odd number and he an even one. So suppose *C* was wrong to say that neither of us can work out who has the greater number.'

The train stops at a small station. 'If I can, it can only be by using the mathematical induction to place *B* with the greater. But, in that case, he can also use it to place me with the greater. So, if *C* is wrong, we can each prove the other to have the greater. That is absurd too. So *C* is right after all.'

Off goes the train again. 'It might follow, perhaps, that whichever of us has the lesser can pin the greater on the other. For instance, if *B* were here, we could have a little dialogue, which went "I do not know who is greater", "Nor I", "Nor I", and so on until we reached the lower number. At that point its owner would say, "I know that yours is greater!". We could even conduct the dialogue in silence, if we had a convention that each second's silence represented a failure to deduce who was greater. But, no, without this source of extra premises, neither of us can know that he is the lesser. So *C* is still right.'

Another rude whistle comes from the engine. 'Of course each of us has to assume that the other is not stupid. But it is not asking much – it would be plain enough to someone who chose *1* that the other number is greater; and this can be banked, in readiness for proving that, if *C* is right, no one chose *2*. So, I fear, I can sum up pretty neatly. If *C* is right, no one chose any number. But I chose *157*. So *C* is wrong. So at least one of us can work out who has the greater number. I cannot. *B* cannot. So *C* is right.'

At this moment the collector of inference tickets interrupts and the train of thought is broken. Readers of this journal are invited to sort the matter out.

University of East Anglia,
Norwich NR4 7TJ

©MARTIN HOLLIS 1984

THERE IS NO SET OF ALL TRUTHS

By PATRICK GRIM

AN important philosophical consequence of Cantor's work has apparently been overlooked. There can be no set of all truths.

I

The proof is as follows.

Suppose that there *is* a set of all truths $\mathscr{T}$:

$$\mathscr{T} = \{T_1, T_2, T_3, \ldots\},$$

and consider further all subsets of $\mathscr{T}$, elements of the power set $\mathscr{P}\mathscr{T}$:

$$\begin{array}{l} \phi \\ \{T_1\} \\ \{T_2\} \\ \{T_3\} \\ \vdots \\ \{T_1, T_2\} \end{array}$$

$\{T_1, T_3\}$

$\vdots$

$\{T_1, T_2, T_3\}$

$\vdots$

Now to each element of this power set will correspond a truth. To each element of the power set, for example, T_1 either will or will not belong as a member. In either case we will have a truth:

$T_1 \notin \phi$
$T_1 \in \{T_1\}$
$T_1 \notin \{T_2\}$
$T_1 \notin \{T_3\}$

$\vdots$

$T_1 \in \{T_1, T_2\}$
$T_1 \in \{T_1, T_3\}$

$\vdots$

$T_1 \in \{T_1, T_2, T_3\}$

$\vdots$

There is of course nothing special about T_1 here – we could have used any particular truth in its place. There are also myriad other ways of constructing a distinct truth for each element of the power set $\mathscr{P}\mathscr{T}$.

To each element of the power set will correspond a distinct truth, and thus there will be at least as many truths as there are elements of the power set $\mathscr{P}\mathscr{T}$. But by Cantor's power set theorem the power set of any set will be *larger* than the original.[1] There will then be *more* truths than there are members of $\mathscr{T}$. Some truths must be left out, and thus $\mathscr{T}$ cannot, as assumed, be a set of *all* truths.

II

Let me mention just one application of the argument above, against a common approach to possible worlds.

Possible worlds are often introduced as maximal consistent sets of propositions – proposition-saturated sets to which no further proposition can be added without precipitating inconsistency – or as some sort of fleshed-out correlates to such sets.[2] The *actual* world, on such an account, is the maximal consistent set of propositions all members of which actually obtain – a maximal and consistent set of all *truths* – or is an appropriately fleshed-out correlate to such a set.

---

[1] See for example Irving M. Copi, *Symbolic Logic*, fifth edition (New York: Macmillan, 1979), pp. 189–90.

[2] See for example Robert Merrihew Adams, 'Theories of Actuality', *Noûs*, 17 (1974), 211–31, and Alvin Plantinga's treatment of worlds in terms of books in *God, Freedom, and Evil* (Grand Rapids, Michigan: Wm. B. Eerdmans, 1980), pp. 35–44, and *The Nature of Necessity* (Oxford: Clarendon Press, 1974), pp. 44–69.

By the argument above, however, there can be no set of all truths. *Any* set of true propositions will leave some true proposition out, and thus there can be no maximal set of truths. Given this notion of possible worlds, then, there can be no *actual* world.[3]

III

The general argument above, of course, applies explicitly only against a *set* of all truths. It quite clearly relies, moreover, on a crucial assumption of bivalence regarding set membership.

We might then hope to dispel the air of paradox and to save a category of all truths by recourse to many-valued set theories or to the non-set *classes* of alternative set theories.

Here let me say simply that I am not sanguine about our prospects. Many-valued logics exhibit many-valued forms of the Liar and of Russell's paradox,[4] and my guess is that they will exhibit many-valued forms of the Cantorian argument above as well. Alternative set theories seem capable of including a universal class only at some unacceptable cost, such as crippling mathematical induction.[5] My guess is that the same may hold for any attempt to include even a *class* of all truths.

It might appear at first glance that there is a conflict between the Cantorian result above and Lindenbaum's Lemma, in terms of which we *can* construct maximal proof-theoretically consistent sets for familiar formal systems.[6] The conflict is merely apparent, however, since (for one thing) Lindenbaum's Lemma relies crucially on the fact that wffs of such systems are explicitly finite. No such limitation is imposed on the truths of $\mathscr{T}$ in the Cantorian argument.

Lindenbaum's Lemma can be seen, however, as preserving a notion of maximal proof-theoretically consistent sets for certain systems, and possible worlds construed in terms of them, as important tools for the logician. The possible worlds that the Cantorian result impugns are those grander entities, corresponding to sets of *all truths*, so tempting to the metaphysician.

State University of New York,
Stony Brook, NY 11794, U.S.A.

©Patrick Grim 1984

---

[3] For a similar argument against such an approach to possible worlds, using a variation on the Liar, see my 'Some Neglected Problems of Omniscience', *American Philosophical Quarterly*, 20 (1983), 265-76.

[4] See esp. Nicholas Rescher, *Many-Valued Logic* (New York: McGraw-Hill, 1969), pp. 87-90, 206-12.

[5] See esp. W. V. O. Quine, *Set Theory and Its Logic* (Cambridge, Mass.: Belknap, Harvard University Press, 1963), pp. 287-389.

[6] See for example Geoffrey Hunter, *Metalogic* (Berkeley: University of California Press, 1971), pp. 110-11 and 177-8, and Elliot Mendelson, *Introduction to Mathematical Logic* (Princeton: D. Van Nostrand, 1964), pp. 64-5 and 93.

Morality and Conflict

STUART HAMPSHIRE

'Characteristically authoritative . . . Stuart Hampshire compels us to think. A single review cannot do justice to its richness, nor indeed to the moving seriousness with which it is written.' *The Listener*
'Important and influential essays' *Times Higher Education Supplement*
186 pages, hardback **£16.50** (0 631 13336 4)
paperback **£6.50** (0 631 13982 X)

Realism and Truth

MICHAEL DEVITT

This book argues that realism is essentially a metaphysical issue. Drawing on the correspondence theory of truth, Michael Devitt develops a strong defence of realism, which leads him to advance a naturalized approach to semantics and epistemology. In the discussion which ensues he examines and rejects in turn the views of most influential critics of realism.
264 pages, **£19.50** (0 631 13535 9)

Women, Reason and Nature

Some Philosophical Problems with Feminism

CAROL McMILLAN

'It is good to have a serious book on feminist thought.'
Ann Oakley, *British Book News*
Carol McMillan argues an original and provocative case: that feminism, far from presenting a radical challenge to sexism, actually shares its most fundamental and mistaken assumptions about women's place in society.
176 pages, hardback **£16.50** (0 631 12496 9)
paperback **£5.95** (0 631 13894 3)

Dialectic of Nihilism

GILLIAN ROSE

This book fundamentally challenges the radical credentials of post-structuralism. Though Derrida, Foucault and Deleuze claim to have 'deconstructed' metaphysics, Gillian Rose shows that their work has much in common with previous attempts to 'end' the metaphysical tradition.
240 pages, hardback **£22.50** (0 631 13191 4)
paperback **£8.50** (0 631 13708 4)

Basil Blackwell

Personal Identity

SYDNEY SHOEMAKER and RICHARD SWINBURNE

What does it mean to say that this person at this time is 'the same' as that person at an earlier time? If the brain is damaged or the memory lost, how far does a person's identity continue? In this book two eminent philosophers develop very different approaches to the problem.

Great Debates in Philosophy
164 pages, hardback **£15.00** (0 631 13208 2)
paperback **£5.50** (0 631 13432 8)

Consciousness and Causality

A Debate on the Nature of Mind

D.M. ARMSTRONG and NORMAN MALCOLM

How are the objects of consciousness related to consciousness? Can consciousness be defined as an awareness of inner mental states. Can one perceive consciousness in oneself? Can states of consciousness be identified as brain states? David Armstrong and Norman Malcolm here present opposing views on these and related questions in the philosophy of mind.

Great Debates in Philosophy
228 pages, hardback **£17.50** (0 631 13212 0)
paperback **£6.50** (0 631 13433 6)

From Knowledge to Wisdom

A Revolution in the Aims and Methods of Science

NICHOLAS MAXWELL

In this wide-ranging yet closely argued book, the author argues forcefully the need for a comprehensive revolution in the aims and methods of science. He urges a new kind of enquiry which would have as its fundamental aim the enhancement of personal and social wisdom.
320 pages, **£19.50** (0 631 13602 9)

Truth, Knowledge and Modality

Philosophical Papers, Volume 3

G.H. von WRIGHT

In this third volume, Professor von Wright concentrates on the ideas of truth, knowledge and modality, and the articles included are possibly more original and more recent than those of the earlier volumes. In subject matter, he is closer here to the topics which concern him most, and which will be of greatest interest to other philosophers.
176 pages, **£17.50** (0 631 13367 4)

Basil Blackwell

Forthcoming in Spring 1985

Volume 12:

Contemplation and Action: 1902-1914

the third of 28 beautifully-bound volumes containing

The Collected Papers of Bertrand Russell

A major contribution to twentieth century scholarship

Advisory Editorial Board:
Sir Alfred Ayer, Ursula Franklin, Philip Gaskell,
Yvon Gauthier, Ivor Grattan-Guinness, Jock Gunn,
H.C.G. Matthew, John Paismore, D.F. Pears,
John M. Robson, Katherine Tait, John W. Yolton.

'The entire series of volumes, when completed, will be indispensable to a thorough study of the intellectual development of one whose influence on the philosophy of his and our time has perhaps been greater than that of any other single individual.' *T.L.S.*

George Allen & Unwin (Publishers) Ltd.,
PO Box 18, Park Lane,
Hemel Hempstead, Herts HP2 4TE.

George Allen & Unwin
London Boston Sydney

MIND

A QUARTERLY REVIEW
OF PHILOSOPHY
edited by D. W. Hamlyn
with the assistance of S D. Guttenplan

Vol XCIII No 372 October 1984

Articles

The Hypothetical Imperative: R. S. DOWNIE
Are 'Scientific' Objects Coloured?: C. L. HARDIN
Does the Gettier Problem Rest on a Mistake?: R. L. KIRKHAM
Mind and Matter: A Problem that Refuses Dissolution: A. SKILLEN
The Contingent Identity of Particulars and Universals: A. CASULLO
High-minded Egoism and the Problem of Priggishness: N. M. LEMOS

Discussions

An Argument against Anti-realist Semantics: A. APPIAH
Thalberg on Inmateriality: P. GLASSEN
Frege, the Proliferation of Force, and Non-cognitivism: S. C. HURLEY
I Will, if You Will: C. RADFORD
Reply to Schueler on Akrasia: C. C. W. TAYLOR
Paul Russell on Hume's 'Reconciling Project': A. FLEW
Understanding Happiness: G. HAGBERG

Critical Notice

Wittgenstein on Rules and Private Language by S. A. Kripke: R. SCRUTON

Book Reviews Books Received Notices

Orders to: Journals Department,
Basil Blackwell,
108 Cowley Road,
Oxford OX4 1JF, England

Annual subscription:
Institutions: £14.50 or $32.50 (Canada $37.75)
Individuals: £10.00 or $22.50 (Canada $26.25)
Single issues: £4.00 or $9.00 (Canada $10.50)

Published for the MIND ASSOCIATION by BASIL BLACKWELL

AUSTRALASIAN JOURNAL OF PHILOSOPHY

Contents of Vol. 62, No. 3; September 1984

EDITOR: BRIAN ELLIS

ARTICLES

David Lewis Putnam's Paradox
Peter J. Markie Feinberg on Moral Rights
Stephen Hetherington Parsons and Possible Objects
Jay E. Harker Can There Be an Infinite Regress of Justified Beliefs
Michael H. Robins Deviant Causal Chains and Non-Basic Action

DISCUSSION

Richard J. Hall Pierre and the New World Makers
Albert Casullo Conjunctive Properties Revisited
Robert Elliot and Andre Gallois Would it Have Been Me? (Against the Necessity of Origin)
Christopher S. Hill Watsonian Freedom and Freedom of the Will

BOOKS RECEIVED, NOTES AND NEWS, INDEX

ANNUAL SUBSCRIPTIONS: Individuals A$16 (US$18 or Stg£9); Institutions A$32, (US$36 or Stg£17). All correspondence should be addressed to the Editor or Business Manager, Department of Philosophy, La Trobe University, Bundoora, Victoria, 3083, Australia.

PUBLISHED QUARTERLY BY
THE AUSTRALASIAN ASSOCIATION OF PHILOSOPHY

Journal for the Theory of Social Behaviour

General Editors: C.W. Smith, P. Collett,
and S. Moscovici

Volume 14 no. 3 October 1984

Contents

Role and Rational Action	MARGARET URBAN COYNE
A Conceptual Analysis of Self-Disclosure	D. V. FISHER
A Critical Evaluation of Altman's Definition of Privacy as a Dialectical Process	W. H. FODDY
Comparative Narratives: Some Rules for the Study of Action	PETER ABELL
Action as a Text: Gadamer's Hermeneutics and the Social Scientific Analysis of Action	SUSAN HEKMAN
Purposes, Conditioning, and Skinner's Moral Theory: Comments on Mill's Observations	BRUCE WALLER
Purpose and Conditioning: A Reply to Waller	JOHN A. MILLS

1984 subscriptions:
Institutions £28.95 (inland) £35.50 (overseas) $74.00 (USA) $90.60 (Canada);
Individuals £12.95 (inland) £15.95 (overseas) $35.95 (USA) $44.00 (Canada)
Single copies are available, please write for prices.

Orders to: JOURNALS DEPARTMENT, **BASIL BLACKWELL,**
108 Cowley Road, Oxford OX4 1JF, England.

The Journal of Philosophy

Subscriptions at $15.00/year; $12.00 to students
$20.00/year to libraries

Students' price extended to retired and unemployed philosophers. Published bimonthly 1904 to 1976, monthly thereafter. Complete volumes and all separate issues available back to January 7, 1904 (volume I, number 1); prices as follows:

BACK VOLUMES (from the first to the penultimate year)
32-page issues, $1.50; 64-page, $3.00; special issues, $4.00.
Volumes (unbound, volume indexes included), $30.00 each.

CURRENT VOLUMES (this year and last)
64-page issues, $2.00; special issues, 96-page or over, $3.00.
Volumes (unbound, volume indexes included), $25.00 each.

ALSO AVAILABLE

Cumulative Fifty-year Index, 1904–1953; articles classified by subject and author; 452 p.; cloth, $12.00. Ten-year Supplement, 1954–1963: 98 p.; cloth $3.00, paper $2.00.

720 PHILOSOPHY HALL, COLUMBIA UNIVERSITY, NYC 10027

Teaching Philosophy

Contents — July, 1984 — Vol. 7/3

Craig Ihara *Moral Skepticism and Tolerance*
Berel Lang *Tolerance and Evil*
David McKenzie *The Fundamentalist Student*
Jeffrey Berger *Writing to Learn in Philosophy*
Robert Brumbaugh *The Mathematical Imagery of Plato,* Republic *X*
Fred Oscanyan *Critical Thinking: Reply to Moore*
Brooke Moore *Critical Reflection: Response to Oscanyan*

Case Studies in the Ethics of Philosophy Teaching

Philip Pecorino *Introduction*
Norman Bowie & Richard Hart *"Should We Be Discussing This?"*
Lisa Newton & Richard Wright *"The Borrowed Syllabus"*

Subscriptions (4 issues): $15 individuals; $29 others. Add $2 for postage outside the USA. Send orders to the Philosophy Documentation Center, Bowling Green State University, Bowling Green, OH 43403.
Ed by Arnold Wilson, Dept of Phil, Univ of Cincinnati, Cincinnati, OH 45221-0206.

METAPHILOSOPHY

Vol. 15 : 3/4 — **Editor: Terrell Ward Bynum** — July/October 1984

Coercion and the authority of reason — JOHN CHURCHILL
Descriptive epistemology — JANE DURAN
A defense of epistemic intuitionism — PAUL K. MOSER
Frege's justificationalism: truth and the recognition of authority — M. A. NOTTURNO
Who needs paradigms? — PAUL A. ROTH
Knowability, believability and begging the question: a reply to Sanford — J. I. BIRO
Comments on a claim that some phenomenological statements may be a posteriori — M. M. VAN DE PITTE
Reply to Professor van de Pitte — RICHARD SCHMITT
Inexhaustibility and ontological plurality — STEPHEN DAVID ROSS
Mutual aid and selfish genes — JANNA L. THOMPSON

1984 Subscription rates
Institutions: £28.00 (UK) £36.50 (Overseas) $65.00 (USA) $79.75 (Canada)
Individuals: £16.50 (UK) £20.20 (Overseas) $38.30 (USA) $47.00 (Canada)
Orders, with payment, to: Sue Dommett, Journals Department, Basil Blackwell, 108 Cowley Road, Oxford, OX4 1JF.

NOTES

The ANALYSIS Committee consists of: Chairman, P. T. Geach; Secretary, T. R. Baldwin; Andrew Harrison, R. H. Holland, J. H. McDowell, D. H. Mellor, M. A. Stewart, R. G. Swinburne, A. R. White, C. J. F. Williams, Peter Winch. This committee is responsible for appointing and advising the Editor and for the general policy of the paper.

SUBSCRIPTIONS. ANALYSIS is published by Basil Blackwell Publisher Ltd, 108 Cowley Road, Oxford OX4 1JF, England. Each volume comprises four numbers, each of at least 48 pages, appearing in January, March, June and October. Orders should be sent to the publisher. Subscription prices are as follows:

	U.K.	Overseas	U.S.A.	Canada
Institutions	£11.00	£13.50	$33.00	$39.50
Individuals	£7.00	£8.60	$21.00	$25.00

U.S. MAILING. Second class postage paid at New York. Postmaster: Send address corrections to: Analysis: c/o Expediters of the Printed Word Ltd, 515 Madison Avenue, New York, NY 10022 (U.S. Mailing Agent).

CONTRIBUTIONS. Articles submitted for publication should be addressed to Mr Christopher Kirwan, Exeter College, Oxford OX1 3DP.

Contributors are asked to note the following requirements. Articles should normally not exceed 3,000 words in length, including references and footnotes. They should be type written in double spacing on one side of the paper only; photographic copies are acceptable. Footnotes should be kept to a minimum and wherever possible avoided altogether. Inset quotations and formulae should be well spaced from the text preceding and following them. The spacing of symbols within formulae should be carefully attended to. Single quotation marks should be used, except for purposes of internal quotation and "scare" quotes. Discussion papers should be sent in as soon as possible after the appearance of the article to which they refer.

Typescripts of rejected articles will be returned only if self-addressed envelopes with correct postage are provided (in the case of overseas contributors, international reply coupons sufficient for airmail).

Proofs of accepted articles will be sent to authors for correction, together with information about free copies. Typescripts of accepted articles will be retained by the Editor on the assumption that authors have kept their own copies.

Apart from correction of printers' errors, only very minor changes in proof will normally be accepted by the Editor, to avoid delay and expense in production.

17 FEB 1985

Typeset by Unicus Graphics Ltd, Horsham. Printed and bound in Great Britain by
Whitstable Litho Ltd, Whitstable